

MOAT-B

MRP 1 ~~HALRGFCSDGSDPLWDWNTWNTSNPDFTKCFONTVLWVPCFYLWACFPYFLYLSRHDRCYIOMTPLNKTALGFLWIVCWADLFYSWERSRG~~ 100

MOAT-B 1 ~~.....~~ MRP 3

MRP 101 ~~FLAPVFLVSPSTLLGTTTATFLIOLEERRKGVOSSGIMLTFWLVALVALCALA1LRSKIMTALKEDAQVQLFRDITFVVYFSLLLQVLVLSFCSDRSPLFSE~~ 200

MOAT-B 4 ~~VYOEVKPNPLQDANICSRVFFWLNPLFKIGHKRRLEEDDMYSVLPEDRSQSOHGEELQFWDKEVLAENDAQK~~ 77

MRP 201 ~~TIHDPNCPCESSASFLSRITFWITGLIVRGYRQPLEGSDLWLSNKEDTSEQVPPVLUKNWKKCEAKTRKOPVKVYSSKDPKAOPKESSKVDANEVEAL~~ 300

MOAT-B 78 ~~PSLTRAIKCYKSYLVLGIFTLIEESAKVIOFLCKIINYFENYDPMDSVALNTAYAYATVLTFCCTLIL.AILHHLYFYHVOCAGNRL~~ 166

MRP 301 ~~IVKSPQKEWNPSPFLKVLYKTFGPYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNNDTAKPDWOGY...FTVLLFVTACLOTLVLHQYFHICFVSGHRI~~ 395

MOAT-B 167 ~~RVAMCHMIYRKALRLSNMAMGKTTGQIVNLLSNVDNKFDOTVFLHFLWAGPLOIAJAVTALLWMEIGISLAGHAVLITILLPLQSCFGKLFSSLRSKTA~~ 266

MRP 396 ~~KTAVICAVYRKALVITNSARKSSTVGEIVNLSMSDAQRFMDLATYINMWSAPLOVILALYLLWLNGLPSVLAGVAVMVLVPMVNAVMANKTKYQVAJM~~ 495

MOAT-B 267 ~~TFTDARIRTMNEVITGIRIICKHYAEWSFSNLSITNLRKKEISKILRSSLRGMNLASFFSASKIIVFVFTTYYVLG..SVITASRUVFVAVTLYGAVRLT~~ 364

MRP 496 ~~KSKDNRIKLMNEILNGIKVLKLYAWELAFKDVKLAIRQEEKLVLLKSAYLSAVGFTTWCTPFLVALCTFAVVYTIIDENNIQDAQTAFVSLALFNLRLFP~~ 595

MOAT-B 365 ~~VTLFFFPSATIEVSEAISSIRIOTFLLDEIS...QRNRQLPSDGKIKMVHQDPTAFWDKASEPTFTLQGLSTVPGELLAVVGPVGACKSSLLSAVLG~~ 460

MRP 596 ~~LNI.LPMVISSIVQASVSLKRLRIFLSHEELEPDPSIERRPVKDCGGTNSITVRNATFTWAR.SDPPTLNGITFSIPEGALVAVVGQVGGKLSLLSALLA~~ 693

MOAT-B 461 ~~ELAPSHGLVSVHRIAYVSSQOPWFSGTLRSNIIIFGKKYEKERYEKVIKACALKDQLLEDGDLTVIGDRGTTLSCGQOKARVNALARAVYQDADIYLLDD~~ 560

MRP 694 ~~EMDKVECHVAIKGSVAVPQOAWIQONSDLSRENILFGCOLEEPYRSVIQACALLPDIEILPSGDRTEIGKVNLSCGQOKORVSLARAVYNSADIYLFDD~~ 793

MOAT-B 561 ~~PLSAVDAEVSRHLFELCICQ..ILHEKITILVTHQLQYLKAASQILILKDGKVMQKGTYTEFLKSGIDFGSLLK.....KNEESEOPPPVPG....~~ 645

MRP 794 ~~PLSAVDAHVGKJIFENVIGPKGMLKNTTRILVTHSMSYLPQVDVIVMSGGKISEMGSYQELLARDGAFELRTYASTEQDAEENGVTGUSCPGKEA~~ 893

MOAT-B 646 ~~...TPTLRRNRTFSESSVWSQSSRPSLKDGALESQDT..ENVPTLSEENRSEGKVGFQAYKNYFRAGAHWIVFIFLILLNTAAOVAYVLQ~~ 731

MRP 894 ~~KOMENGMLVTDAGKQLOLSSSSYSCDISRHNNSTAELOKAECKEETWKLMEADKAQTCQVKLSVWYDYMKAIGLFLSIFLF.NCNHVSALAS~~ 992

MOAT-B 732 ~~DWLSYWANKQSHLNVTVNGGNVTEKLDNNYLGIYSGLTVATVFLGVIARSLLVFLVNVSSOTLHNHOFESTLKAQVLFDRNPIGRILNRFSKDIGH~~ 831

MRP 993 ~~NYWLSLWTD....DPIVNGTQEHTKVR....LSVYGAQGISQGIAVFGYMSAVSIGGILASRCLHVLDLHSILRSRPMSSFTPSGNLNVRFSKELDT~~ 1082

MOAT-B 832 ~~LDDLLPLTFDFTQTLQVQCVSVAVAVIPWIAIPVPLVPLGIIFIFLRRYFLETSRDKRLESTTRSPVFHSLLLQGLWTIRAYKAEERCQELFDAHO~~ 931

MRP 1083 ~~VDSMIPVIMKFMGSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQLKLESVSRSPVSHNETLLGVSVIRAFEEQERFIHQSDLKV~~ 1182

MOAT-B 932 ~~DHSEAWFLFLTTSRWFVAFRLDAICAMFVIIVAFGLSLILAKTLDAGQVGLALSYALTLMGFOWCVROSAEVENMISVERVIEYTDLEKEAPWEYOK.R~~ 1030

MRP 1183 ~~DENOKAYYPSIVANRMLAVRLECVCNCIVLFAALFASRHSLSAGLGLVLSVSYSLQVTTYNLWLRVMSSEMETNIVAVERKEYSETEKEAPWQIQTFR~~ 1282

MOAT-B 1031 ~~PPPAWPHEGVIFDNVNFHYSPPGPPVLUKHLTALIKSOEKVGIVGRTGAGKSSLISALFRlse..PEGKIIWIDKILTEICLHDLRKKMSIIPOEPVLFtg~~ 1129

MRP 1283 ~~PPSSWQVGRVEFRNYCLRYREDLDFVLRHINVTINGEKGIVGRTGAGKSSLTLGLFRINESAECEIIIDGINIAKIGLHDLRFKTYIIIPDPVLFSG~~ 1382

MOAT-B 1130 ~~TMRKKNLDPFKEMTDEELWNALQEVOLKETIEDLPGKMDTEAEGSNFSVGOROLVCLARAILRKNOIIIDEATANVDPRTDELIQKKIREKFAHCTVL~~ 1229

MRP 1383 ~~SLRMNLDPFSQYSDEEWVTSLELAHLKDFVSALPDKLDHECAEGGENLSVQRLVCLARALLRKTKILVLDEATAAVDLETDOLI0STIRTOFEDCTVL~~ 1482

MOAT-B 1230 ~~TIARLNTIIDSDKIMVLDSGRLKEYDEPVYLLQNKESLYKVMQQLGKAEAAALTETAKQVYFKRNYHIGHTDMVNTNSGOPSTLTIFETAL~~ 1325

MRP 1483 ~~TIARLNTIMDYTRVIVLKDGEI0EYGAPSDLQQR.GLFYMAKDAGLV~~ 1531

Figure 1

2/56

Fig. 2A

Nucleotide Binding Fold I									
H.MOAT-B	WILSFTVNG	COLLAVVCP	GACKSSLLS	VEGETADPSH	15	...	S	HGRIAVVSQ	PRD-ESTLRS
S.YCF1	KNTNEIAKK	ENCTCINVK	GSNTAILLSC	MCGDERRDK	FR	...	70	HGSVAVVSQ	PRHNNNTVKE
H.MRP	ENCTERKIFP	ENAVVWQ	GACKSSLLSA	LNADMKV	HE	...	71	KESVAVVPA	AVQNDLSN
H.CMOAT	VRDVNLGTHA	SKKAVVIFP	GACKSSLLSA	MGLMENVIG	KV	...	72	KITTAVYPLS	SQGQNTIKD
H.CFTR	ENKNTENR	COLLAVVST	GACKSSLLMM	IMMELPSH	KY	...	73	SRISFCSQ	SQGQNTTKE
L.PgpA	ENKNTENR	SKKAVVIFP	GACKSSLLSA	LNDEYSEVS	PQ	...	74	ENSAVAVPQ	AGMMNLNG
H.SUR	ENKNTENR	ENAVVWQ	GACKSSLLLA	ALGHMKVSH	PA	...	75	ENSAVAVPQ	ENMMNLNG
H.MDR1	ENKNTENR	ENAVVWQ	GACKSSLLSA	MGRKQPSH	PA	...	76	ENSAVAVPQ	PRMMNLNG
							77	ENSAVAVPQ	PRMMNLNG
							78	ENSAVAVPQ	PRMMNLNG
							79	ENSAVAVPQ	PRMMNLNG
							80	ENSAVAVPQ	PRMMNLNG
							81	ENSAVAVPQ	PRMMNLNG
							82	ENSAVAVPQ	PRMMNLNG
							83	ENSAVAVPQ	PRMMNLNG
							84	ENSAVAVPQ	PRMMNLNG
							85	ENSAVAVPQ	PRMMNLNG
							86	ENSAVAVPQ	PRMMNLNG
							87	ENSAVAVPQ	PRMMNLNG
							88	ENSAVAVPQ	PRMMNLNG
							89	ENSAVAVPQ	PRMMNLNG
							90	ENSAVAVPQ	PRMMNLNG
							91	ENSAVAVPQ	PRMMNLNG
							92	ENSAVAVPQ	PRMMNLNG
							93	ENSAVAVPQ	PRMMNLNG
							94	ENSAVAVPQ	PRMMNLNG
							95	ENSAVAVPQ	PRMMNLNG
							96	ENSAVAVPQ	PRMMNLNG
							97	ENSAVAVPQ	PRMMNLNG
							98	ENSAVAVPQ	PRMMNLNG
							99	ENSAVAVPQ	PRMMNLNG
							100	ENSAVAVPQ	PRMMNLNG
							101	ENSAVAVPQ	PRMMNLNG
							102	ENSAVAVPQ	PRMMNLNG
							103	ENSAVAVPQ	PRMMNLNG
							104	ENSAVAVPQ	PRMMNLNG
							105	ENSAVAVPQ	PRMMNLNG
							106	ENSAVAVPQ	PRMMNLNG
							107	ENSAVAVPQ	PRMMNLNG
							108	ENSAVAVPQ	PRMMNLNG
							109	ENSAVAVPQ	PRMMNLNG
							110	ENSAVAVPQ	PRMMNLNG
							111	ENSAVAVPQ	PRMMNLNG
							112	ENSAVAVPQ	PRMMNLNG
							113	ENSAVAVPQ	PRMMNLNG
							114	ENSAVAVPQ	PRMMNLNG
							115	ENSAVAVPQ	PRMMNLNG
							116	ENSAVAVPQ	PRMMNLNG
							117	ENSAVAVPQ	PRMMNLNG
							118	ENSAVAVPQ	PRMMNLNG
							119	ENSAVAVPQ	PRMMNLNG
							120	ENSAVAVPQ	PRMMNLNG
							121	ENSAVAVPQ	PRMMNLNG
							122	ENSAVAVPQ	PRMMNLNG
							123	ENSAVAVPQ	PRMMNLNG
							124	ENSAVAVPQ	PRMMNLNG
							125	ENSAVAVPQ	PRMMNLNG
							126	ENSAVAVPQ	PRMMNLNG
							127	ENSAVAVPQ	PRMMNLNG
							128	ENSAVAVPQ	PRMMNLNG
							129	ENSAVAVPQ	PRMMNLNG
							130	ENSAVAVPQ	PRMMNLNG
							131	ENSAVAVPQ	PRMMNLNG
							132	ENSAVAVPQ	PRMMNLNG
							133	ENSAVAVPQ	PRMMNLNG
							134	ENSAVAVPQ	PRMMNLNG
							135	ENSAVAVPQ	PRMMNLNG
							136	ENSAVAVPQ	PRMMNLNG
							137	ENSAVAVPQ	PRMMNLNG
							138	ENSAVAVPQ	PRMMNLNG
							139	ENSAVAVPQ	PRMMNLNG
							140	ENSAVAVPQ	PRMMNLNG
							141	ENSAVAVPQ	PRMMNLNG
							142	ENSAVAVPQ	PRMMNLNG
							143	ENSAVAVPQ	PRMMNLNG
							144	ENSAVAVPQ	PRMMNLNG
							145	ENSAVAVPQ	PRMMNLNG
							146	ENSAVAVPQ	PRMMNLNG
							147	ENSAVAVPQ	PRMMNLNG
							148	ENSAVAVPQ	PRMMNLNG
							149	ENSAVAVPQ	PRMMNLNG
							150	ENSAVAVPQ	PRMMNLNG
							151	ENSAVAVPQ	PRMMNLNG
							152	ENSAVAVPQ	PRMMNLNG
							153	ENSAVAVPQ	PRMMNLNG
							154	ENSAVAVPQ	PRMMNLNG
							155	ENSAVAVPQ	PRMMNLNG
							156	ENSAVAVPQ	PRMMNLNG
							157	ENSAVAVPQ	PRMMNLNG
							158	ENSAVAVPQ	PRMMNLNG
							159	ENSAVAVPQ	PRMMNLNG
							160	ENSAVAVPQ	PRMMNLNG
							161	ENSAVAVPQ	PRMMNLNG
							162	ENSAVAVPQ	PRMMNLNG
							163	ENSAVAVPQ	PRMMNLNG
							164	ENSAVAVPQ	PRMMNLNG
							165	ENSAVAVPQ	PRMMNLNG
							166	ENSAVAVPQ	PRMMNLNG
							167	ENSAVAVPQ	PRMMNLNG
							168	ENSAVAVPQ	PRMMNLNG
							169	ENSAVAVPQ	PRMMNLNG
							170	ENSAVAVPQ	PRMMNLNG
							171	ENSAVAVPQ	PRMMNLNG
							172	ENSAVAVPQ	PRMMNLNG
							173	ENSAVAVPQ	PRMMNLNG
							174	ENSAVAVPQ	PRMMNLNG
							175	ENSAVAVPQ	PRMMNLNG
							176	ENSAVAVPQ	PRMMNLNG
							177	ENSAVAVPQ	PRMMNLNG
							178	ENSAVAVPQ	PRMMNLNG
							179	ENSAVAVPQ	PRMMNLNG
							180	ENSAVAVPQ	PRMMNLNG
							181	ENSAVAVPQ	PRMMNLNG
							182	ENSAVAVPQ	PRMMNLNG
							183	ENSAVAVPQ	PRMMNLNG
							184	ENSAVAVPQ	PRMMNLNG
							185	ENSAVAVPQ	PRMMNLNG
							186	ENSAVAVPQ	PRMMNLNG
							187	ENSAVAVPQ	PRMMNLNG
							188	ENSAVAVPQ	PRMMNLNG
							189	ENSAVAVPQ	PRMMNLNG
							190	ENSAVAVPQ	PRMMNLNG
							191	ENSAVAVPQ	PRMMNLNG
							192	ENSAVAVPQ	PRMMNLNG
							193	ENSAVAVPQ	PRMMNLNG
							194	ENSAVAVPQ	PRMMNLNG
							195	ENSAVAVPQ	PRMMNLNG
							196	ENSAVAVPQ	PRMMNLNG
							197	ENSAVAVPQ	PRMMNLNG
							198	ENSAVAVPQ	PRMMNLNG
							199	ENSAVAVPQ	PRMMNLNG
							200	ENSAVAVPQ	PRMMNLNG
							201	ENSAVAVPQ	PRMMNLNG
							202	ENSAVAVPQ	PRMMNLNG
							203	ENSAVAVPQ	PRMMNLNG
							204	ENSAVAVPQ	PRMMNLNG
							205	ENSAVAVPQ	PRMMNLNG
							206	ENSAVAVPQ	PRMMNLNG
							207	ENSAVAVPQ	PRMMNLNG
							208	ENSAVAVPQ	PRMMNLNG
							209	ENSAVAVPQ	PRMMNLNG
							210	ENSAVAVPQ	PRMMNLNG
							211	ENSAVAVPQ	PRMMNLNG
							212	ENSAVAVPQ	PRMMNLNG
							213	ENSAVAVPQ	PRMMNLNG
							214	ENSAVAVPQ	PRMMNLNG
							215	ENSAVAVPQ	PRMMNLNG
							216	ENSAVAVPQ	PRMMNLNG
							217	ENSAVAVPQ	PRMMNLNG
							218	ENSAVAVPQ	PRMMNLNG
							219	ENSAVAVPQ	PRMMNLNG
							220	ENSAVAVPQ	PRMMNLNG
							221	ENSAVAVPQ	PRMMNLNG
							222	ENSAVAVPQ	PRMMNLNG
							223	ENSAVAVPQ	PRMMNLNG
							224	ENSAVAVPQ	PRMMNLNG
							225	ENSAVAVPQ	PRMMNLNG
							226	ENSAVAVPQ	PRMMNLNG
							227	ENSAVAVPQ	PRMMNLNG
							228	ENSAVAVPQ	PRMMNLNG
							229	ENSAVAVPQ	PRMMNLNG
							230	ENSAVAVPQ	PRMMNLNG
							231	ENSAVAVPQ	PRMMNLNG
							232	ENSAVAVPQ	PRMMNLNG
							233	ENSAVAVPQ	PRMMNLNG
							234	ENSAVAVPQ	PRMMNLNG
							235	ENSAVAVPQ	PRMMNLNG
							236	ENSAVAVPQ	PRMMNLNG
							237	ENSAVAVPQ	PRMMNLNG
							238	ENSAVAVPQ	PRMMNLNG
							239	ENSAVAVPQ	PRMMNLNG
							240	ENSAVAVPQ	PRMMNLNG
							241	ENSAVAVPQ	PRMMNLNG
							242	ENSAVAVPQ	PRMMNLNG
							243	ENSAVAVPQ	PRMMNLNG
							244	ENSAVAVPQ	PRMMNLNG
							245	ENSAVAVPQ	PRMMNLNG
			</						

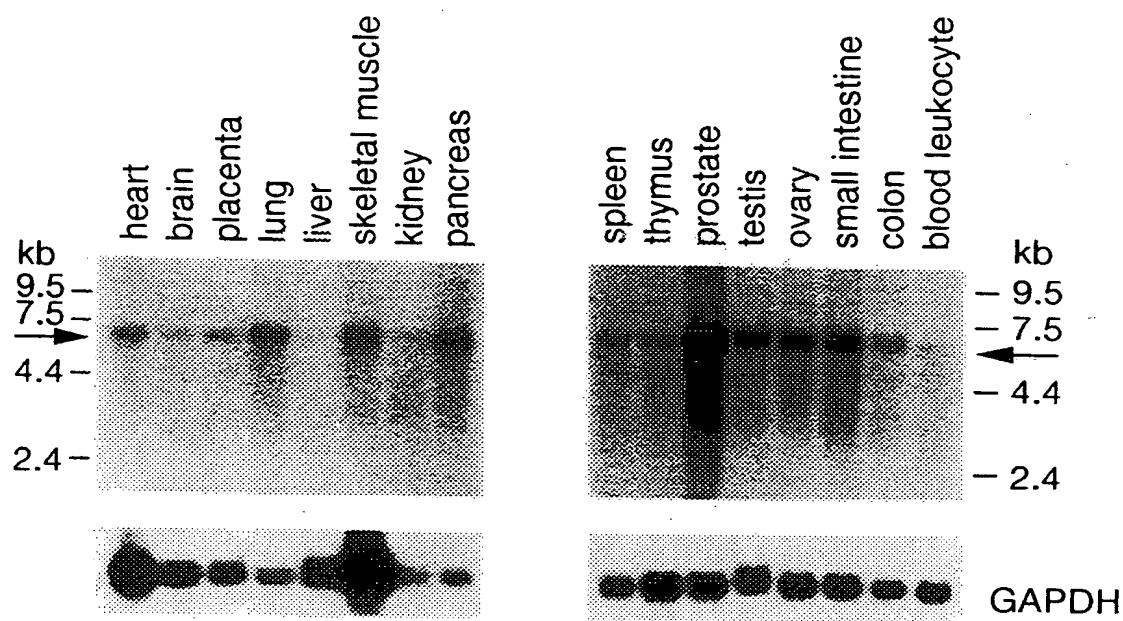


Figure 3

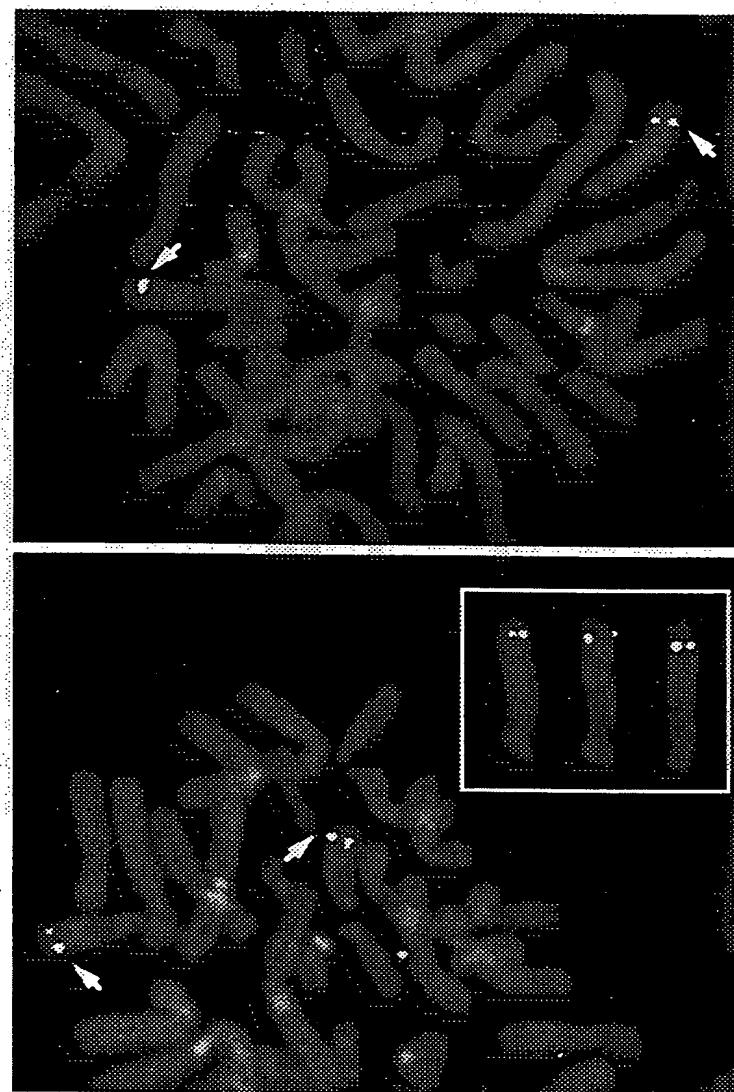


Figure 4

Fig. 5A

1 MKDIDIGKEY IIPSPGYRSV RERTSTSGTH RDREDSKFRR TRPLECQDAL ETAARAEGLS  
 61 LDASMHSQLR ILDEEHPKGK YHHGLSALKP IRTTSKHQHP VDNAGLFSCM TFSWLSSLAR  
 121 VAHKKGELSM EDVWSLSKHE SSDVNCRRLE RLWQEELNEV GPDAASLRRV VWIFCRTRLI  
 TM1  
 181 LSIVCLMITQ LAGFSGPAFM VKHLLEYTQA TESNLQYSLL LVLGLLLTEI VRSWSLALTW  
 TM2  
 241 ALNYRTGVRL RGAILTMASK KILKLNKIE KSLGELINIC SNDGQRMFEA AAVGSLLAGG  
 TM3  
 301 PVVAILGMIY NVIILGPTGF LGSAVFILFY PAMMFASRLT AYFRRKCVAA TDERVQKMNE  
 TM4  
 361 VLTYIKFIKM YAWVKAFSQS VQKIREEERR ILEKAGYFOQ ITVGVAPIVV VIASVVTFSV  
 TM5  
 421 HMTLGFDLTA AQAFTVVTVF NSMTFALKVT PFSVKSLSEA SVAVDRFKSL FILMEEVHMIK  
 481 NKPASPHIKI EMKNATLAWD SSHSSIQNSP KLTPKMKDK RASRGKKEKV RQLORTEHQ  
 541 VLAEQKGHLL LDSDERPSPE EEEGKHIHLG HRLQRTLHS IDLEIQEGKL VGICGSVGSG  
 TM6  
 601 KTSLISAILG QMTLLEGSIA ISGTFAYVAQ QAWILNATLR DNILFGKEYD EERYNSVLS  
 661 CCLRPDLAIL PSSDLTEIGE RGANLSGGQR QRISLARALY SDRSIYILDD PLSALDAHVG  
 NBF1 C B  
 721 NHIFNSAIRK HLKSKTVLV THQLQYLVDC DEVIFMKEGC ITERGTHEEL MNLNGDYATI  
 781 FNNLLLGETP PVEINSKKET SGSGQKKSQDK GPKTGSVKKE KAVKPEEGQL VQLEEKQGS  
 TM7  
 841 VPWSVYGVYI QAAGGPLAFL VIMHALFMLNV GSTAFSTWWL SYWIQQSGN TTVTRGNETS  
 TM8  
 901 VSDSMKDNPB MQYIASIYAL SMAVMLILKA IRGVVFVKGTRASSRLHDE LFRRIILRSPM  
 961 KFDDTTPTGR ILNRFSKDMQ EVDVRLPFQA EMFIQNVILV FFCVGMIAVG FPWFLVAVGP  
 TM9  
 1021 LVILFSVLHI VSRVLIRELK RLDNITQSPP LSHITSSIQG LATIHAYNKG QEFHRYQEL  
 TM10  
 1081 LDDNQAPFFL FTCAMRWLAV RLDLISIALI TTTGLMIVLM HQQIPPAYAG LAISYAVQLT  
 TM11  
 1141 GLFQFTVRLA SETEARFTSV ERINHYIKTL SLEAPARIKN KAPSPDWPQE GEVTFENAEM  
 NBF2  
 1201 RYRENPLPLV KKVSFTIKPK EKIGIVGRTG SGKSSLGMAF FRLVELSGGC IKIDGVRISD  
 A  
 1261 IGLADLRSKL SIIPQEPVLF SGTVRSNLD PNOYTEDQIW DALERTHMKE CIAQLPLKLE  
 1321 SEVMENGDNF SVGERQLLCI ARAILRHCKI LILDEATAAM DTETDILLIQE TIREAFADCT  
 TM12  
 1381 MLTIAHRLHT VLGSDRIMVL AQGQVVEFDT PSVLLSNDS RFYAMFAAAE NKVAVKG

Fig. 5B

1 MGPMDALCGS GELGSKFWDS NLSVHTENPD TM1 LTPCFQNSLL AWVPCIYLWV ALPCYLLYLR  
 61 HHCRGYIILS HLSKLKMVLG TM2 VLLWCWSWAD LFYSFHGLVH GRAPAPVFFF TPLVVGVTL  
 121 LATLLIQQYER LQGVQSSGVL TM3 IIFWFLCVVC AIVPFRSKIL LAKAEGEISD PFRFTTFYIH  
 181 FALVLSALIL ACFREKPPFF SAKNVDPNPY PETSVGFLSR LFFWWFTKMA IYGYRHPLEE  
 241 KDLWSLKEED RSQMVVQQLL EAWRKQEKT ARHKASAAPG KNASGEDEVL LGARPRPRKP  
 301 SFLKALLATF GSSFLISACF KLIQDLLSFI TM4 NPQLLSILIR FISNPMAPSW TM5 WGFLVAGLMF  
 361 LCSMMQSLIL QHYTHYIFVT GVKFRTGIMG VIYRKALVIT NSVKRASTVG EIVNLMSVDA  
 421 QRFMDLAPFL NLLWSAPLQI ILAIYFLWQN TM6 LGPSVLAGVA FMVLLIPLNG AVAVKMRAFQ  
 481 VKQMQLKDSR ICLMSEILNG IKVLKLYAWE PSFLKQVEGI RQGELQLLRT AAYLHTTTF  
 541 TM10 TWMCSPPFLVT LITLWVYVYV DPNNVLDAAK AFVSVSLFNI LRLPLNMLPQ TM11 LISNLTQASV  
 601 SLKRIQQFLS QEELDPQSVE RKTISPGYAI TIESGTFTWA QDLPPTLHSL  $\rightarrow$  NBF1 DIOQVPKGALV  
 661 AVVGPVCCGK SSLVSALLGE MEKLEGKVHM KGSVAYVPOQ AWIONCTLQE NVLFGKALNP  
 721 A KRYQQTLEAC ALLADLEMP GGDQTEIGEK GINLSSGORQ RVSLARAVYS DADIFLDDP  
 781 LSAVDSEVAK HIFDHEVIGPE GVLAGKTRVL VTHGISFLPQ TDFIIVLADG QVSEMPYPA  
 841 LLQRNGSFAN FLCNYAPDED QGHLEDSWTA LEGAEDKEAL LIEDTLSNHT DLTDNDPVY  
 901 VVQKQFMQQL SALSSDGEQ GRPVPRRHLG PSEKVQVTEA KADGALTQEE KAAIGTVELS  
 961 VFWDYAKAVG LCTTIAICLL YVGQAAAIG ANWWSANTN DAMADSRQNN  $\bullet$  TSLRLGVYAA  
 1021 TM12 LGILQGFLVM LAAMAMAAGG IQAARVLHQI LLHNKIRSPQ SFFDTTPSGR TM13 ILNCFSKDIY  
 1081 VVDEVLAPEI LMLLNSFFNA ISTLWVIMAS TM14 TPLFTVVLIP LAVLYTLVQR FYAATSRQLK  
 1141 RLESVSRSPY YSHFSETVTG ASVIRAYNRS RDFEIIISDTK VDANQRSCYP YIISNRWLSI  
 1201 TM16 GVEFVGNCVV LFAALFAVIG RSSLNPGLVG LSVSYSILQVT FALNWHIRMM SDLESNIVAV  
 1261 ERVKEYSKTE TEAPWVVEGS RPPEGWPPRG EVEFRNYSVR YRPGLDLVLR DLSLHVHGGE  
 1321 KVGIVGRTGA GKSSMTLCLF RILEAKGEI RIDGLNVADI GLHDLRSQLT IIPQDPILFS  
 1381 A GTLRMNLDPP GSYSEEDIWW ALELSHLHTF VSSQPAGLDF QCSEGGENLS VGORQLVCLA  
 1441 RALLRKSRIL TM17 VLDEATAAID LETDNLIQAT IRTQFDCTV LTIAHRLNTI MDYTRVLVLD  
 1501 B KGVVAEFDSP ANLIAARGIF YGMARDAGLA C

**Nucleotide Binding Fold I**

MOAT-D	PSHSDTQVPK	GAIVAVVGVV	CGGKSSRVER	LQFEVEKIEG	KF	HM	KPSVAYVPGC	AKGSCCLC	707
MRP	NGVTTSPPE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
cMOAT	VRDVNLDTMA	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
MOAT-C	PSHSDTQVQE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
MOAT-B	PSHSDTQVRP	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
CFTR	KLKDNFKFEE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
SUR	LSNTIRPR	GCUTMPSV	CGGKSSSLSA	ALQENQKVSS	AFWSSLPDS	EIGEDPSPER	ETATDLDI	DIRK	785
MDR1	KGGLNLKVOS	PSHSDTQVLP	CGGKSSSLSA	MORLYDPMPS	MPSVVDGQDIR	TINVRFLREI	IGV	PE	486

A

MOAT-D	NYLFCQKA	LN	PKYVQOTILEN	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MRP	NYLFCQLE	LE	PSYVRSVQCA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
cMOAT	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MOAT-C	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MOAT-B	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
CFTR	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
SUR	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MDR1	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793

B

MOAT-D	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MRP	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
cMOAT	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MOAT-C	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MOAT-B	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
CFTR	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
SUR	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793
MDR1	NYLFCQLE	FN	PKYVQAVLEA	CAILKPLM	EGCGCPEIGE	KGINLSSGQK	QFVSLARAVV	SLADYLQD	PLSAVOSVVA	KHIDHV	793

C

MOAT-D	PSHSDTQVPK	GAIVAVVGVV	CGGKSSRVER	LQFEVEKIEG	KF	HM	KPSVAYVPGC	AKGSCCLC	707
MRP	NGVTTSPPE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
cMOAT	VRDVNLDTMA	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
MOAT-C	PSHSDTQVQE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
MOAT-B	PSHSDTQVRP	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
CFTR	KLKDNFKFEE	GAIVAVVGVV	CGGKSSSLSA	LQAEKDKVEG	HV	AI	KPSVAYVPGC	AKGSCCLC	724
SUR	LSNTIRPR	GCUTMPSV	CGGKSSSLSA	ALQENQKVSS	AFWSSLPDS	EIGEDPSPER	ETATDLDI	DIRK	785
MDR1	KGGLNLKVOS	PSHSDTQVLP	CGGKSSSLSA	MORLYDPMPS	MPSVVDGQDIR	TINVRFLREI	IGV	PE	486

**Nucleotide Binding Fold II**

MOAT-D	IKDLSLHVHO	GEKVGIVGVRT	GAHKSSMTLC	LFPIISAAAG	PLRIGGLNVR	DPLGLHDRLRQ	ITIIPCDPFI	PSGIPML	DPFGS	SE	1392		
MRP	LRHINVING	GEKVGIVGVRT	GAHKSSLTLC	LFPIISAAAG	EITIDGGINER	KIGLHDRLRQ	ITIIPCDPFI	PSGSILRM	DPFGS	SE	1396		
cMOAT	LRGITCDEGS	MEX1GIVGVRT	GAHKSSLTNC	LFPIISAAAG	OMIICGVDIA	SIGLHDRLRQ	ITIIPCDPFI	PSGSILRM	DPFGS	SE	1403		
MOAT-C	IKKVSFTKPK	MEX1GIVGVRT	GAHKSSLGMG	LFPLVLSCS	CSKIDVDRVS	DIGLADLRSK	LSLIPCDPFI	PSGIVPSNL	DPFGS	SE	1296		
MOAT-B	IKHKLALIKS	GEKVGIVGVRT	GAHKSSLISA	LFPLVLSC	PSG	KWID	ILTT	EKMLDRLRQK	MSLIPCDPFI	PSGIVPSNL	DPFGS	SE	1143
CFTR	IKHKLALISP	GEKVGIVGVRT	GAHKSSFSLA	FEPMVDTFEG	HIIIDGIDIA	KLFQHTRKSS	LSLIPCDPFI	PSGIVPSNL	PSGIVPSNL	DPFGS	SE	1447	
SUR	LENISFSDSP	GEKVGIVGVRT	GAHKSSFLSA	FEPMVDTFEG	HIIIDGIDIA	KLFQHTRKSS	LSLIPCDPFI	PSGIVPSNL	PSGIVPSNL	DPFGS	SE	1312	
MDR1	IQGSLLEVKK	PSHSDTQVLP	GAHKSTVVQL	FEFYDPLA	KVLLKESKIK	RLNVQWBAH	IGIVSIEFIS	EDCSIAE	IA	YQNSRVVQ	Q	1142	

A

MOAT-D	EDPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	ASRFLVLD	DEA	TAANQLE	DT	1465
MRP	EDWTSLELA	HLKDNISALE	DLGKHCAGC	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	DT	TAANQLE	DT	1469
cMOAT	EEIIVKALBLA	HLKDNISALE	DLGKHCAGC	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	DT	TAANQLE	DT	1476
MOAT-C	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1369
MOAT-B	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1216
CFTR	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1385
SUR	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1520
MDR1	SEIVRVAKEA	NIHATIESLS	NGYSTKVGDR	CTCGECSK	PLCLARALLS	STKILVDEA	TAANQLE	OPHILLE	SEI	ST	1215

C

MOAT-D	EDPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	ASRFLVLD	DEA	TAANQLE	DT	1465
MRP	EDWTSLELA	HLKDNISALE	DLGKHCAGC	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	DT	TAANQLE	DT	1469
cMOAT	EEIIVKALBLA	HLKDNISALE	DLGKHCAGC	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	DT	TAANQLE	DT	1476
MOAT-C	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1369
MOAT-B	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1216
CFTR	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1385
SUR	DEPQWALELS	SHHTFSSCP	ACGDFQCS	EG	GENLSVGDRG	PLCLARALLS	STKILVDEA	TAANQLE	TAANQLE	DT	1520
MDR1	SEIVRVAKEA	NIHATIESLS	NGYSTKVGDR	CTCGECSK	PLCLARALLS	STKILVDEA	TAANQLE	OPHILLE	SEI	ST	1215

B

Fig. 6A

8/56

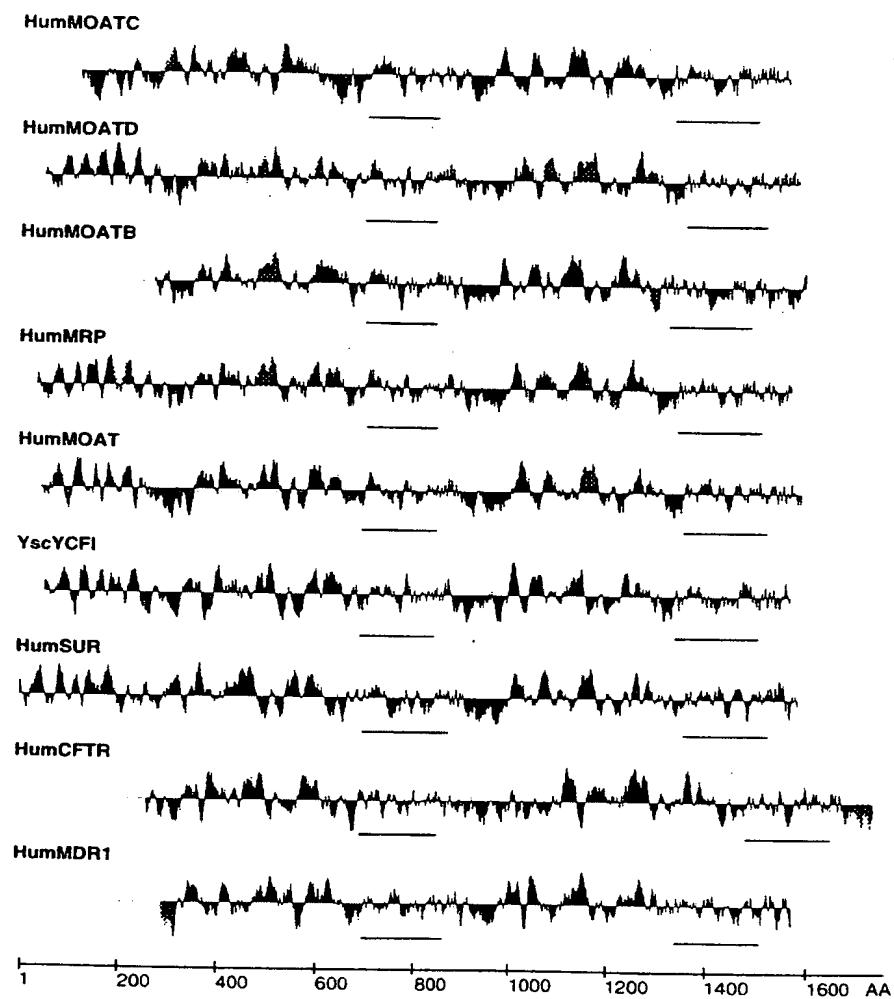


Fig. 6B

9/56

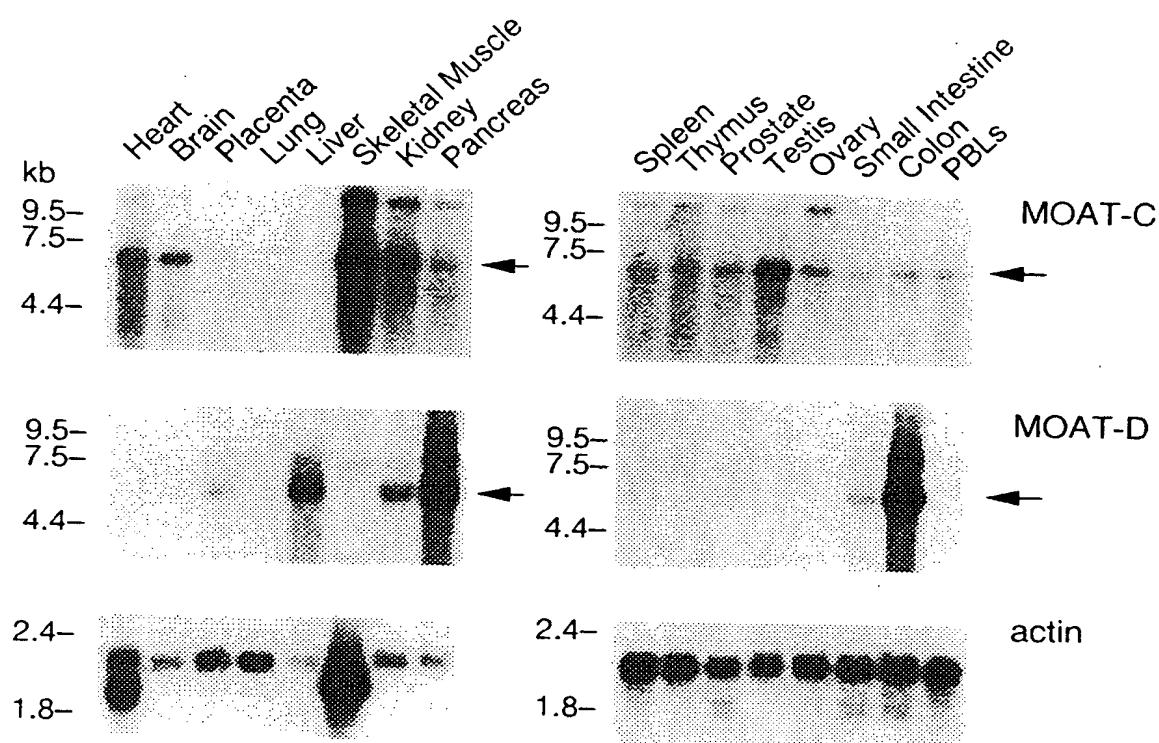


Figure 7

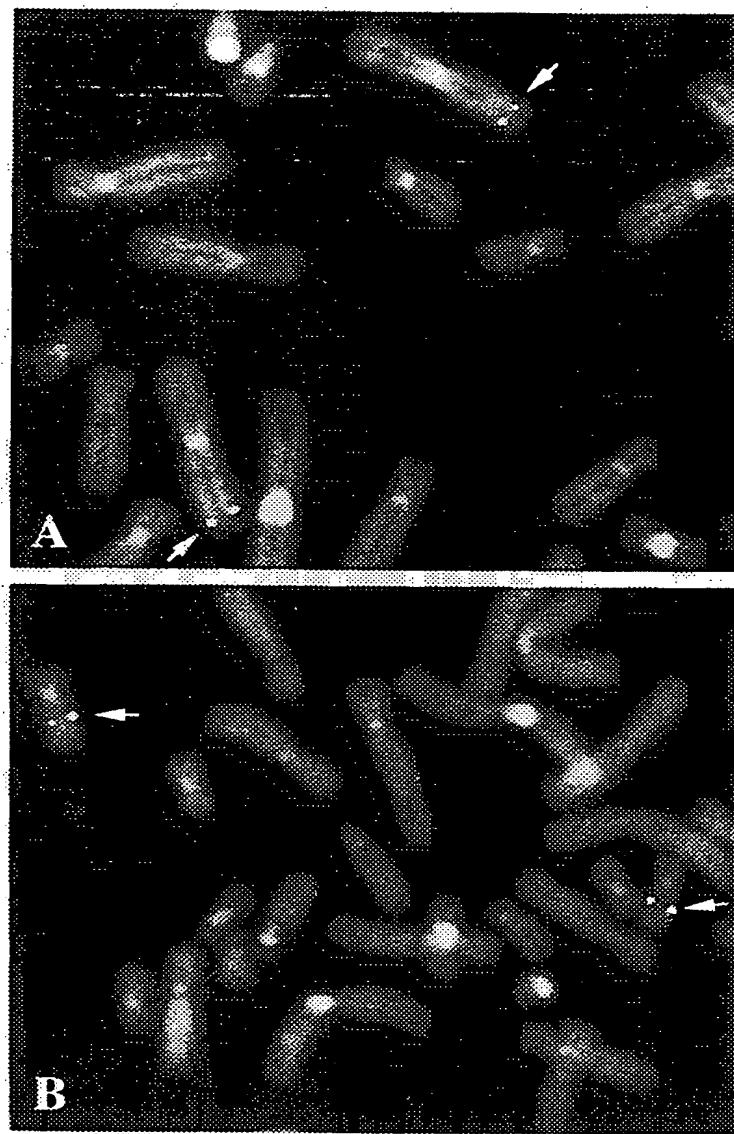


Figure 8

1 MAAPAEPCAG QGVWNQTEPE PAATSLSLC FLRTAGVWVP PMYLWVLGPI YLLFIHHHGR  
 61 GYL RMSPLFK AKMVLGFALI VLCTSSVAVA LWKIQQGTPE APEFLIHPTV WLTTMSFAVF  
 121 LIH TERKKGV QSSGVLFYIW LLCFVLPATN AAQOASGAGF QSDPVRHLST YLCLSLVVAQ  
 181 FVL SCLADQP PFFPEDPQQS NPCPETGAAF PSKATFWWVS GLVWRGYRRP LRPKDLWSLG  
 241 RENSSEELVS RLEKEWMRNR SAARRHNKAI AFKRKGGSQM KAPETEPFLR QEGSQWRPLL  
 301 KAIWQVFHST FLLGTLSLII SDVFRFTVPK LLSLFLEFIG DPKPPANKGY LLAVLMFLSA  
 361 CLQTLPEQQN MYRLKVPQMR LRSAITGLVY RKVLALSSGS RKASAVGDVV NLVSVDVQRL  
 421 TESVLYLNGL WLPLVWIVVC FVYLVQLLGP SALTAIAVFL SLLPLNFFIS KKRNHHQEEQ  
 481 MRQKDSRRL TSSILRNSKT IKFHGWEGAF LDRVIGIRQ ELGALRTSGL LFSVSLVSFQ  
 541 VSTFLVALVV FAVHTLVAEN AMNAEKAFVT LTVLNILNKA QAFLPFSIHS LVQARVSFDR  
 601 LVTFLCLEEV DPGVVVDSSSS GSAAGKDCIT IHSATFAWSQ ESPPCLHRIN LTVPOGCLLA  
 661 VVGPVGAGKS SLLSALLGEL SKVEGFVSIE GAVAYVPQEA WVQNTSVVEN VCFGQELDPP  
 721 WLERVLEACA LQPDVDSFPE GIHTSIGEQG MNLSGGOKQR LSLARAVYRK AAVYLLDDPL  
 781 AALDAHVGQH VFNQVIGPGG LLQGTTRILV THALHILPQA DWIIVLNGA IAEMGSYQEL  
 841 LQRKGALVCL LDQARQPGDR GEGETEPGTS TKDPRGTSAG RRPELRRERS IKSVPEKDRT  
 901 TSEAQTEVPL DDPDRAGWPA GKDSIQYGRV KATVHLAYLR AVGTPCLLYA LFLFLCQQVA  
 961 SFCRGYWLWLSL WADDPAVGGQ QTQALRGGI FGLLGCLQAI GLFASMAAVL LGGARASRLL  
 1021 FQRLLWDVVR SPISFFERTP IGHLLNRFSK ETDTVDVDIP DKLRSLLMYA FGLLEVSLLV  
 1081 AVATPLATVA ILPLFLYAG FQSLYVVSSC QLRRLESASY SSVC SHMAET FQGSTVVRAF  
 1141 RTOAPFVAQN NARVDESQRI SFPRLVADRW LAANVELLGN GLVFAATCA VLSKAHLSAG  
 1201 LVGF SVSAAL QVTQALQWVV RNWTDLENSI VSVERMQDYA WTPKEAPWRL PTCAAQPPWP  
 1261 QGGQIEFRDF GLRYRPELPL AVQGVSLKIH AGEKVGIVGR TGAGKSSLAS GLLRLQEAAE  
 1321 GGIWIDGVPI AHVGLHTLRS RISIIIPQDPI LFPGSLRMNL DLLQEHSDEA IWAALETVQL  
 1381 KALV ASLPGQ LOYK CADRGE DLSVGQKQLL CLAR ALLRKT QILILDEATA AVDPGTELOM  
 1441 QAMLG SWFAQ CTVLLIAHRL RSVMDCARVL VMDKGQVAES GSPAQLLAQK GLFYRLAQES  
 1501 GLV

Figure 9

12/56

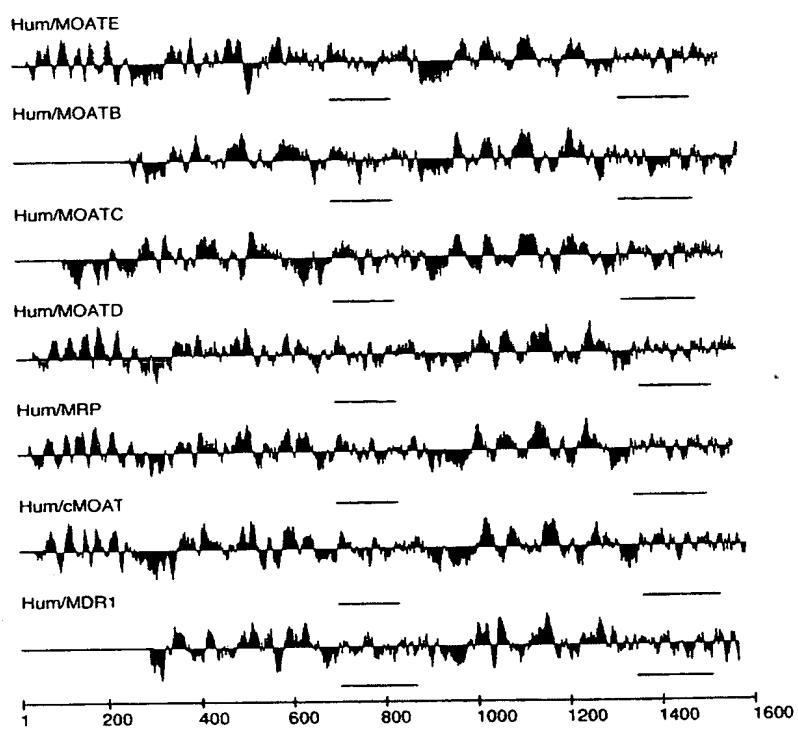
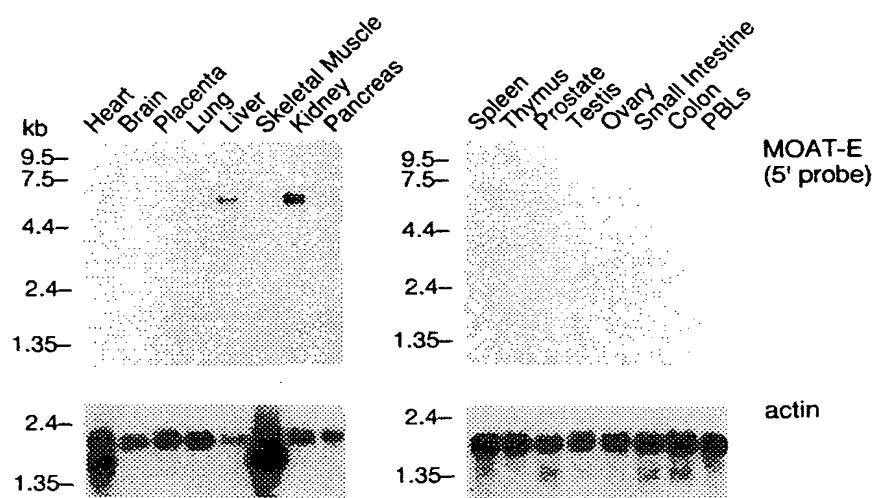


Figure 10



**Figure 11**

## MOAT B cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGCTGCCGTGTACCAAGGAGGTGAAGCCCAACCGCTGCAGGACGCGAACATCTGCTCA  
 1 ----- + ----- + ----- + ----- + ----- + 60  
 TACGACGGGCACATGGTCCTCCACTTCGGGTTGGCGACGTCTGCCTGCGCTTAGACGAGT

a M L P V Y Q E V K P N P L Q D A N I C S -

CGCGTGTCTCTGGTGGCTCAATCCCTGTTAAAATTGGCCATAACGGAGATTAGAG  
 61 ----- + ----- + ----- + ----- + ----- + 120  
 GCGCACAAAGAACCGACCGAGTTAGGGAACAAATTAAACCGGTATTCGCTCTAAC

a R V F F W W L N P L F K I G H K R R L E -

GAAGATGATATGATTCACTGCTGCCAGAACCGCTCACAGCACCTGGAGAGGGAGTTG  
 121 ----- + ----- + ----- + ----- + ----- + 180  
 CTTCTACTATACATAAGTCACGACGGTCTCTGGCGAGTGTGCGTGGAACCTCTCCTAAC

a E D D M Y S V L P E D R S Q H L G E E L -

CAAGGGTTCTGGATAAAAGAAGTTTAAGAGCTGAGAACGACAGAACGCTCTTTA  
 181 ----- + ----- + ----- + ----- + ----- + 240  
 GTTCCCAAGACCCATTCTCAAAATTCTGACTCTACTGCGTGTCTCGGAAGAAAT

a Q G F W D K E V L R A E N D A Q K P S L -

ACAAGAGCAATCATAAAGTGTACTGGAAATCTTATTAGTTGGAAATTACGTTA  
 241 ----- + ----- + ----- + ----- + ----- + 300  
 TGTTCTCGTTAGTATTCACAATGACCTTACAATAATCAAAACCTTAAATGCAAT

a T R A I I K C Y W K S Y L V L G I F T L -

ATTGAGGAAAGTGCCAAAGTAATCCAGCCCATTGGAAAAATTATTAATTATTT  
 301 ----- + ----- + ----- + ----- + ----- + 360  
 TAACTCCTTCACGGTTCTAGGTGGTATAAAAACCTTTAATAATTAATAAAAA

a I E E S A K V I Q P I F L G K I I N Y F -

GAAAATTATGATCCCATGGATTCTGTGGCTTGAACACAGCGTACGCCATGCCACGGTG

Figure 12A

361 -----+-----+-----+-----+-----+ 420  
 CTTTAATACTAGGGTACCTAAGACACCGAAACTGTGTCGATGCGATAACGGTCCAC

a E N Y D P M D S V A L N T A Y A Y A T V -

CTGACTTTTGACGCTCATTGGCTATACTGCATCACTTATATTTATCACGTTCA  
 421 -----+-----+-----+-----+-----+ 480  
 GACTGAAAAACGTGCGAGTAAACCGATATGACGTAGTGAATATAAAAATAGTCAAGTC

a L T F C T L I L A I L H H L Y F Y H V Q -

TGTGCTGGGATGAGGTTACGAGTAGCCATGTGCCATATGATTATCGGAAGGCACTTGT  
 481 -----+-----+-----+-----+-----+ 540  
 ACACGACCCCTACTCCAATGCTCATCGGTACACGGTATACTAAATAGCCTCCGTGAAGCA

a C A G M R L R V A M C H M I Y R K A L R -

CTTAGTAACATGGCCATGGGAAGACAACACAGGCCAGATAGTCATCTGCTGTCCAAT  
 541 -----+-----+-----+-----+-----+ 600  
 GAATCATTGTACCGGTACCCCTCTGTTGGTGTCCGGTCTATCAGTTAGACGACAGGTTA

a L S N M A M G K T T T G Q I V N L L S N -

GATGTGAACAAGTTGATCAGGTGACAGTGTCTTACACTTCTGTGGCAGGACACTG  
 601 -----+-----+-----+-----+-----+ 660  
 CTACACTTGTCAAAGTAGTCCACTGTACAAGAATGTGAAGGACACCCGTCCTGGTGAC

a D V N K F D Q V T V F L H F L W A G P L -

CAGGCGATCGCAGTGAUTGCCCTACTCTGGATGGAGATAGGAATATCGTCCTGCTGG  
 661 -----+-----+-----+-----+-----+ 720  
 GTCCGCTAGCGTCACTGACGGATGAGACCTACCTCTATCCTTATAGCACGGAACGACCC

a Q A I A V T A L L W M E I G I S C L A G -

ATGGCAGTTCTAATCATTCTCTGCCCTTGCAAAGCTGTTGGGAAGTTGTTCTCATCA  
 721 -----+-----+-----+-----+-----+ 780  
 TACCGTCAAGATTAGTAAGAGGGACGGAACGTTGACAAAAACCCCTCAACAAGAGTAGT

a M A V L I I L L P L Q S C F G K L F S S -

CTGAGGAGTAAACTGCAACTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATA  
 781 -----+-----+-----+-----+-----+ 840

**Figure 12B**

16/56

GAACCTCATTGACGTTGAAAGTGCCTACGGTCTAGTCCTGGTACTTACCTCAATAT

a L R S K T A T F T D A R I R T M N E V I .

ACTGGTATAAGGATAATAAAAATGTACGCCTGGGAAAAGTCATTTCAAATCTTATTACC  
841 -----+-----+-----+-----+-----+ 900  
TGACCATATTCTATTATTTACATGCGGACCTTTCAAGTAAAAGTTAGAATAATGG

a T G I R I I K M Y A W E K S F S N L I T .

AATTTGAGAAAGAAGGAGATTCCAAGATTCTGAGAAGTTCTGCCTCAGGGGGATGAAT  
901 -----+-----+-----+-----+-----+ 960  
TTAAACTCTTCTTCTAAAGGTTCAAGACTCTCAAGGACGGAGTCCCCTACTTA

a N L R K K E I S K I L R S S C L R G M N .

TTGGCTTCGTTTCACTGCAAGCAAAATCATCGTGTGACCTTCACCACTACGTG  
961 -----+-----+-----+-----+-----+ 1020  
AACCGAAGCAAAAGTCACGTTGTTAGTAGCACAAACACTGGAAGTGGTGGATGCAC

a L A S F F S A S K I I V F V T F T T Y V .

CTCCTCGGCAGTGTGATCACAGGCCAGCCCGTGTGGCAGTGACGCTGTATGGGGCT  
1021 -----+-----+-----+-----+-----+ 1080  
GAGGAGCCGTACACTAGTGTGGTCGGCGACAAGCACCGTCACTGCGACATAACCCGA

a L L G S V I T A S R V F V A V T L Y G A .

GTGCGGCTGACGGTTACCTCTTCCCTCAGCCATTGAGAGGGTGTAGAGGAATC  
1081 -----+-----+-----+-----+-----+ 1140  
CACGCCGACTGCCAATGGAGAAGAAGGGAGTCGGTAACTCTCCCACAGTCTCCGTTAG

a V R L T V T L F F P S A I E R V S E A I .

GTCAGCATCCGAAGAACATCCAGACCTTTGCTACTTGATGAGATATCACAGCGAACCGT  
1141 -----+-----+-----+-----+-----+ 1200  
CAGTCGTAGGCTTCTAGGTCTGGAAAAACGATGAACACTCTATAGTGTGCGTTGGCA

a V S I R R I Q T F L L L D E I S Q R N R .

CAGCTGCCGTCAGATGGTAAAAAGATGGTGCATGTGCAGGATTTACTGCTTTGGGAT  
1201 -----+-----+-----+-----+-----+ 1260  
GTCGACGGCAGTCTACCATTTCTACCACGTACACGTCTAAATGACGAAAAACCTA

Figure 12C

17/56

a Q L P S D G K K M V H V Q D F T A F W D .

AAGGCATCAGAGACCCCAACTCTACAAGGCCTTCCTTACTGTCAGACCTGGCGAATTG  
 1261 -----+-----+-----+-----+-----+-----+ 1320  
 TTCCGTAGTCTCTGGGTTGAGATGTTCCGGAAAGGAAATGACAGTCTGGACCGCTTAAC

a K A S E T P T L Q G L S F T V R P G E L .

TTAGCTGTGGTCGGCCCCGTGGGAGCAGGGAAAGTCATCACTGTTAAGTGCCGTGCTCGGG  
 1321 -----+-----+-----+-----+-----+-----+ 1380  
 AATCGACACCAGCCGGGCACCCCTCGTCCCTCAGTAGTGACAATTACGGCACGAGCCC

a L A V V G P V G A G K S S L L S A V L G .

GAATTGGCCCCAAGTCACGGGCTGGTCAGCGTCATGGAAGAATTGCCTATGTGTCTCAG  
 1381 -----+-----+-----+-----+-----+-----+ 1440  
 CTTAACCGGGGTTCACTGCCCCACCGAGTCGACGTACCTTCTTAACGGATAACAGAGTC

a E L A P S H G L V S V H G R I A Y V S Q .

CAGCCCTGGGTGTTCTCGGAACCTCTGAGGAGTAATATTTATTTGGGAAGAAATATGAA  
 1441 -----+-----+-----+-----+-----+-----+ 1500  
 GTCGGGACCCACAAGAGCCCTTGAGACTCCTCATTATAAAATAACCCCTCTTATACTT

a Q P W V F S G T L R S N I L F G K K Y E .

AAGGAACGATATGAAAAAGTCATAAAGGCTTGTGCTCTGAAAAAGGATTTACAGCTGTTG  
 1501 -----+-----+-----+-----+-----+-----+ 1560  
 TTCCCTTGCTATACTTTTCAGTATTCGAACACGAGACTTTCTAAATGTCGACAAC

a K E R Y E K V I K A C A L K K D L Q L L .

GAGGATGGTATCTGACTGTGATAGGAGATGGGGAAACACGCTGAGTGGAGGGCAGAAA  
 1561 -----+-----+-----+-----+-----+-----+ 1620  
 CTCCTTACCACTAGACTGACACTATCCTCTAGCCCTTGGTGCAGTCACCTCCGTCTT

a E D G D L T V I G D R G T T L S G G Q K .

GCACGGGTAAACCTTGCAAGAGCAGTGTATCAAGATGCTGACATCTATCTGGACGAT  
 1621 -----+-----+-----+-----+-----+-----+ 1680  
 CGTCCCCATTGGAACGTTCTCGTCACATAGTTCTACGACTGTAGATAGAGGACCTGCTA

Figure 12D

18/56

a A R V N L A R A V Y Q D A D I Y L L D D -

CCTCTCAGTGCAGTAGATGCGGAAGTTAGCAGACACTTGTCAACTGTGTATTGTCAA  
 1681 -----+-----+-----+-----+-----+-----+ 1740  
 GGAGAGTCACGTCATCTACGCCCTCAATCGTCTGTGAACAAGCTTGACACATAAACAGTT

a P L S A V D A E V S R H L F E L C I C O -

ATTTTGATGAGAAGATCACAATTTAGTACTCATCAGITGCAGTACCTCAAAGCTGCA  
 1741 -----+-----+-----+-----+-----+ 1800  
 TAAAACGTACTCTCTAGTGTAAAATCACTGAGTAGTCAACGTCATGGAGTTCGACGT

a I L H E K I T I L V T H Q L Q Y L K A A -

AGTCAGATTCTGATATTGAAAGATGGTAAAATGGTGCAGAAGGGGACTTACACTGAGTTC  
 1801 -----+-----+-----+-----+-----+ 1860  
 TCAGTCTAAGACTATAACTTCTACCATTACCGTCTCCCTGAATGTGACTCAAG

a S Q I L I L K D G K M V Q K G T Y T E F -

CTAAAATCTGGTATAGATTTGGCTCCCTTAAAGAAGGATAATGAGGAAAGTGAACAA  
 1861 -----+-----+-----+-----+-----+ 1920  
 GATTTTAGACCATATCTAAAACCGAGGGAAAATTCTTCCTTACTCCTTCACTTGT

a L K S G I D F G S L L K K D N E E S E Q -

CCTCCAGTTCCAGGAACCTCCACACTAAGGAATCGTACCTCTCAGAGTCTCGGTTGG  
 1921 -----+-----+-----+-----+-----+ 1980  
 GGAGGTCAAGGTCTTGAGGGTGTGATTCTTAGCATGGAAGAGTCTCAGAAGCCAAACC

a P P V P G T P T L R N R T F S E S S V W -

TCTCAACAATCTCTAGACCCCTCTTGAAGATGGTGTCTGGAGAGCCAAGATAACAGAG  
 1981 -----+-----+-----+-----+-----+ 2040  
 AGAGTTGTTAGAAGATCTGGGAGGAACCTTCTACCACGAGACCTCTCGGTTATGTCTC

a S Q Q S S R P S L K D G A L E S Q D T E -

AATGTCCCAGTTACACTATCAGAGGAGAACGTTCTGAAGGAAAAGTTGGTTTCAGGCC  
 2041 -----+-----+-----+-----+-----+ 2100  
 TTACAGGGTCAATGTGATAGTCTCCTCTGGCAAGACTTCCTTTCAACCAAAGTCCGG

a N V P V T L S E E N R S E G K V G F Q A

Figure 12E

TATAAGAATTACTCAGAGCTGGTGCCTACTGGATTGTCTTCATTTCTTATTCTCCTA  
 2101 -----+-----+-----+-----+-----+-----+ 2160  
 ATATTCTTAATGAAGTCTCGACCACGAGTGACCTAACAGAAGTAAAAGGAATAAGAGGAT

a Y K N Y F R A G A H W I V F I F L I L L -

AACACTGCAGCTCAGGTTGCCTATGTGCTTCAGATTGGGGCTTCATACTGGGCAAAC  
 2161 -----+-----+-----+-----+-----+-----+ 2220  
 TTGTGACGTCGAGTCCAACGGATAACCGAAGTTCTAACCCACCGAAAGTATGACCCGTTG

a N T A A Q V A Y V L Q D W W L S Y W A N -

AAACAAAGTATGCTAAATGTCACTGTAATGGAGGGAGGAATGTAACCGAGAAGCTAGAT  
 2221 -----+-----+-----+-----+-----+-----+ 2280  
 TTGTTTACAGATTACAGTGACATTACCTCCTCCTTACATTGGCTTCTCGATCTA

a K Q S M L N V T V N G G G N V T E K L D -

CTTAACTGGTACTTAGAATTTCAGGTTAACTGTAGCTACCGTTCTTGGCATA  
 2281 -----+-----+-----+-----+-----+-----+ 2340  
 GAATTGACCATGAATCCTTAAATAAGTCAAATTGACATCGATGGCAAGAAAAACCGTAT

a L N W Y L G I Y S G L T V A T V L F G I -

GCAAGATCTTATTGGTATTCTACGTCTTGTAACTCTTCACAAACTTGCACAACAAA  
 2341 -----+-----+-----+-----+-----+-----+ 2400  
 CGTTCTAGAGATAACCATAAGATGCAGGAACAATTGAGAAGTGTGAAACGTGTTGTT

a A R S L L V F Y V L V N S S Q T L H N K -

ATGTTTGAGTCAATTCTGAAAGCTCCGGTATTATTCTTGATAGAAATCCAATAGGAAGA  
 2401 -----+-----+-----+-----+-----+-----+ 2460  
 TACAAACTCAGTTAAGACTTCGAGGCCATAATAAGAAACTATCTTAGGTTATCCTTCT

a M F E S I L K A P V L F F D R N P I G R -

ATTTAAATCGTTCTCAAAGACATTGGACACTGGATGATTGCTGCCGCTGACGTT  
 2461 -----+-----+-----+-----+-----+-----+ 2520  
 TAAAATTTAGCAAAGAGGTTCTGTAAACCTGTGAACCTACTAAACGACGGGACTGCAA

a I L N R F S K D I G H L D D L L P L T F

Figure 12F

20/56

TTAGATTCATCCAGACATTGCTACAAGTGGTGGTGGTCTCTGTGGCTGTGGCGTG  
 2521 +-----+-----+-----+-----+-----+ 2580  
 AATCTAAAGTAGGTCTGTAACGATGTTACCAACCACACCAGAGACACCGACACCGGAC

a L D F I Q T L L Q V V G V V S V A V A V .

ATTCCTTGGATCGCAATACCCCTGGTCCCGGAAATCATTTCATTTCTTCGGCGA  
 2581 +-----+-----+-----+-----+-----+ 2640  
 TAAGGAACCTAGCGTTATGGGAACCAAGGGAACCTAGTAAAAGTAAAAGAAGCCGCT

a I P W I A I P I V P L G I I F I F L R R .

TATTTTTGGAAACGTCAAGAGATGTGAAGCGCCTGGAATCTACAACCTCGGAGTCCAGTG  
 2641 +-----+-----+-----+-----+-----+ 2700  
 ATAAAAAAACCTTGCAGTTCTCTACACTTCGCGGACCTAGATGTTGAGCCTCAGGTAC

a Y F L E T S R D V K R L E S T T R S P V .

TTTCCCACTTGTCATCTCTCCAGGGCTCTGGACCATCCGGCATACAAAGCAGAA  
 2701 +-----+-----+-----+-----+-----+ 2760  
 AAAAGGGTGAACAGTAGAAGAGAGGTCCCCGAGACCTGGTAGGCCGTATGTTCGTCTT

a F S H L S S S L Q G L W T I R A Y K A E .

GAGAGGTGTCAGGAACGTGGATGCACACCAAGGATTACATTAGAGGCTTGGTTCTG  
 2761 +-----+-----+-----+-----+-----+ 2820  
 CTCTCCACAGTCCTTGACAAACTACGTGTGGTCTAAATGTAAGTCTCCGAACCAAGAAC

a E R C Q E L F D A H Q D L H S E A W F L .

TTTTGACAACGTCCCGCTGGTTCGCCGTCCGTGGATGCCATCTGTGCCATGTTGTC  
 2821 +-----+-----+-----+-----+-----+ 2880  
 AAAAAGTTCAGGGCGACCAAGCGGCAGGCAGACCTACGGTAGACACGGTACAAACAG

a F L T T S R W F A V R L D A I C A M F V .

ATCATCGTTGCCCTGGGTCCCTGATTCTGGCAAAACTCTGGATGCCGGCAGGTTGGT  
 2881 +-----+-----+-----+-----+-----+ 2940  
 TAGTAGCAACGGAAACCCAGGGACTAAGACCGTTTGAGACCTACGGCCGTCCAACCA

a I I V A F G S L I L A K T L D A G Q V G .

TTGGCACTGTCCTATGCCCTCACGCTCATGGGATTTAGTGGTGTTCGACAAAGT

## Figure 12G

2941 + + + + + 3000  
AACCGTGACAGGATACGGGAGTGCAGTACCCCTACAAAGTACCAACACAAGCTGTTCA

a L A L S Y A L T L M G M F Q W C V R Q S .

```

GCTGAAGTTGAGAATATGATGATCTCAGTAGAAAGGGTCATTGAATACACAGACCTTGAAC
3001 ----- + ----- + ----- + ----- + ----- + 3060
CGACTTCAACTCTTATACTACTAGAGTCATTTCCAGTAACCTATGTGTCCTGGAACTT

```

a A E V E N M M I S V E R V I E Y T D I E .

AAAGAACCTGGAAATATCAGAAACGCCACCAACCAGCCTGGCCCCATGAAGGAGTG  
 3061 — + — + — + — + — + — + 3120  
 TTTCTCGTGGAACCCCTATAGTCTTGGGGTGGTCGGACCGGGGTACTTCCCTAC

a K E A P W E Y Q K R P P P A W P H E G V

ATAATCTTGACAATGTGAACCTCATGTACAGTCCAGGTGGGCCTCTGGTACTGAAGCAT  
 3121 — + — + — + — + — + 3180  
 TATTAGAAACTGTTACACTTGAAGTACATGTCAGGTCCACCCGGAGACCATGACTTCGTA

a I I F D N V N F M Y S P G G P L V L K H -

3181 CTGACAGCACTCATTAATCACAAGAAAAGGTTGGCATTGTGGGAAGAACCGGAGCTGGA  
 3240 GACTGTCGTGAGTAATTAGTGTCTTTCCAACCGTAACACCCCTCTGGCCTCGACCT

<sup>a</sup> LTALIKSQEKVGIVGBTGAG -

AAAAGTTCCCTCATCTCAGCCCTTTAGATTGTCAGAACCCGAAGGTAAAATTGGATT  
 3241 — + — + — + — + — + 3300  
 TTTCAAGGGAGTAGAGTCGGGAAAAATCTAACAGTCTGGGCTTCCATTTAAACCTAA

a K S S L I S A L F R L S E P E G K I W I -

3301 GATAAGATCTTGACAACGTAAATTGGACTTCACGATTAAAGGAAGAAAATGTCAATCATA  
 3360 CTATTCTAGAACTGTTGACTTTAACCTGAAGTGCTAAATTCTTCTTTACAGTTAGTAT

**a** D K I L T T E I G L H R I B K K M S I I

3361 — + — + — + — + — + — + 3420

Figure 12H

GGAGTCCTGGACAAAACAAGTGACCTGTTACTCCTTTGGACCTAGGGAAATTCCCTC

a P Q E P V L F T G T M R K N L D P F K E -

CACACGGATGAGGAACGTGGAATGCCTTACAAGAGGTACAACCTAAAGAAACCATTGAA  
3421 -----+-----+-----+-----+-----+-----+ 3480  
GTGTGCCTACTCCTTGACACCTTACGGAATGTTCTCATGTTGAATTCTTGGTAACCT

a H T D E E L W N A L Q E V O L K E T I E -

GATCTTCCTGGTAAAATGGATACTGAATTAGCAGAACATCAGGATCCAATTTAGTGTTGGA  
3481 -----+-----+-----+-----+-----+-----+ 3540  
CTAGAAGGACCATTACCTATGACTTAATCGTCTAGTCCTAGGTTAAAATCACACCT

a D L P G K M D T E L A E S G S N F S V G -

CAAAGACAACGGTGTGCCTGCCAGGGCAATTCTCAGGAAAAATCAGATATTGATTATT  
3541 -----+-----+-----+-----+-----+-----+ 3600  
GTTTCTGTTGACCACACGGAACGGTCCCGTTAAGAGTCCTTTAGTCTATAACTAATAA

a Q R Q L V C L A R A I L R K N Q I L I I -

GATGAAGCGACGGCAAATGTGGATCCAAGAACTGATGAGTTAACACAAAAAAATCCGG  
3601 -----+-----+-----+-----+-----+-----+ 3660  
CTACTTCGCTGCCGTTACACCTAGGTTCTGACTACTCAATTATGTTTTTTAGGCC

a D E A T A N V D P R T D E L I Q K K I R -

GAGAAATTGCCCACTGCACCGTGCTAACCATTCGACACAGATTGAACACCATTATTGAC  
3661 -----+-----+-----+-----+-----+-----+ 3720  
CTCTTAAACGGGTGACGTGGCACGATTGGTAACGTGTCTAACCTGTGGTAATAACTG

a E K F A H C T V L T I A H R L N T I I D -

AGCGACAAGATAATGGTTTAGATTAGGAAGACTGAAAGAATATGATGAGCCGTATGTT  
3721 -----+-----+-----+-----+-----+-----+ 3780  
TCGCTGTTCTATTACCAAAATCTAAGTCCTCTGACTTTCTTATACTACTCGGCATACAA

a S D K I M V L D S G R L K E Y D E P Y V -

TTGCTGAAAATAAGAGAGCCTATTTACAAGATGGTCAACAACTGGCAAGGCAGAA  
3781 -----+-----+-----+-----+-----+-----+ 3840  
AACGACGTTTATTCCTCGGATAAAATGTTCTACCACGTTGTTGACCCGTTCCGTCTT

Figure 12I

23/56

a L L Q N K E S L F Y K M V Q Q L G K A E .  
GCCGCTGCCCTCACTGAAACAGCAAAACAGGTATACTCAAAAGAAATTATCCACATATT  
3841 -----+-----+-----+-----+-----+-----+ 3900  
CGGCGACGGGAGTGACTTTGTCGTTTGTCCATATGAAGTTTCTTAATAGGTGTATAA

a A A A L T E T A K Q V Y F K R N Y P H I .  
GGTCACACTGACCACATGGTTACAAACACTTCCAATGGACAGCCCTGACCTTAACATT  
3901 -----+-----+-----+-----+-----+-----+ 3960  
CCAGTGTGACTGGTGTACCAATGTTGTGAAGGTTACCTGTCGGGAGCTGGAATTGATAA

a G H T D H M V T N T S N G Q P S T L T I .  
TTCGAGACAGCACTG  
3961 -----+--- 3975  
AAGCTCTGTCGTGAC

a F E T A L -

Figure 12J

## MOAT C cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGAAGGATATCGACATAGGAAAAGAGTATATCATCCCCAGTCCTGGGTATAGAAGTGTG  
 1 -----+-----+-----+-----+-----+-----+ 60  
 TACTTCCTATAGCTGTATCCTTTCTCATATAGTAGGGTCAGGACCCATATCTCACAC

a M K D I D I G K E Y I I P S P G Y R S V -

AGGGAGAGAACCGACCTCTGGACGCACAGAGACCGTGAAGATTCCAAGTTCAGGAGA  
 61 -----+-----+-----+-----+-----+ 120  
 TCCCTCTTGGTCGTGAAGACCCCTCGCTGTCTGGCACTTCTAAGGTTCAAGTCTCT

a R E R T S T S G T H R D R E D S K F R R -

ACTCGACCGTTGGAATGCCAAGATGCCCTGGAAACAGCAGCCGAGCCGAGGGCTCT  
 121 -----+-----+-----+-----+-----+ 180  
 TGAGCTGGCAACCTTACGGTTCTACGGAACCTTGTCTCGTCGGCTCGGCTCCGGAGAGA

a T R P L E C Q D A L E T A A R A E G L S -

CTTGATGCCTCCATGCATTCTCAGCTCAGAACATCCTGGATGAGGAGCATCCAAAGGGAAAG  
 181 -----+-----+-----+-----+-----+ 240  
 GAACTACGGAGGTACGTAAGAGTCGAGTCTTAGGACCTACTCCTCGTAGGGTCCCTTC

a L D A S M H S Q L R I L D E E H P K G K -

TACCATCATGGCTTGAGTGCTCTGAAGCCATCCGGACTACTCCAAACACCAGCACCA  
 241 -----+-----+-----+-----+-----+ 300  
 ATGGTAGTACCGAACTCACGAGACTTCGGTAGGCCTGATGAAGGTTGTGGCTGGGT

a Y H H G L S A L K P I R T T S K H Q H P -

GTGGACAATGCTGGCTTTCTGTATGACTTTCTGGCTTCTCTGGCCCGT  
 301 -----+-----+-----+-----+-----+ 360  
 CACCTGTTACGACCCGAAAAAAGGACATACTGAAAAAGCACCAGAAAGAGACCGGGCA

a V D N A G L F S C M T F S W L S S L A R -

GTGGCCCACAAGAAGGGGGAGCTCTCAATGGAAGACGTGTGGTCTGTCCAAGCACGAG

Figure 13A

361 +-----+-----+-----+-----+-----+-----+ 420  
 CACCGGGTGTCTCCCTCGAGAGTTACCTCTGCACACCAGAGACAGGTTCGTGCTC

a V A H K K G E L S M E D V W S L S K H E .  
 TCTTCTGACGTGAAC TGAGAAGACTAGAGAGACTGTGGCAAGAAGAGCTGAATGAAGTT  
 421 +-----+-----+-----+-----+-----+-----+ 480  
 AGAAGACTGCACTTGACGTCTCTGATCTCTGACACCGTTCTCTGACTTACTTCAA

a S S D V N C R R L E R L W Q E E L N E V .  
 GGGCCAGACGCTGCTCCCTCGAAGGGTTGTGGATCTCTGCCGCACCAGGCTCATC  
 481 +-----+-----+-----+-----+-----+-----+ 540  
 CCCGGTCTGCGACGAAGGGACGCTTCCCAACACACACCTAGAACAGACGGCGTGGTCCGAGTAG

a G P D A A S L R R V V W I F C R T R L I .  
 CTGTCCATCGTGTGCCTGATGATCACGCAGCTGGCTGGCTTCAGTGGACCAGCCTTCATG  
 541 +-----+-----+-----+-----+-----+-----+ 600  
 GACAGGTTAGCACACGGACTACTAGTGCCTGACCGACCGAACGTCACCTGGTCGGAAGTAC

a L S I V C L M I T Q L A G F S G P A F M .  
 GTGAAACACCTCTGGAGTATACCCAGGCAACAGAGTCTAACCTGCAGTACAGCTTGTG  
 601 +-----+-----+-----+-----+-----+-----+ 660  
 CACTTTGTGGAGAACCTCATATGGGTCCGTTGTCTCAGATTGGACGTATGTCGAACAAAC

a V K H L L E Y T Q A T E S N L Q Y S L L .  
 TTAGTGTGGCCTCCCTGACGGAAATCGTCGGTCTGGTCGCTTGCAC TGACTTGG  
 661 +-----+-----+-----+-----+-----+-----+ 720  
 AATCACGACCCGGAGGGAGGACTGCCTTAGCACGCCAGAACCGAACGTGACTGAACC

a L V L G L L L T E I V R S W S L A L T W .  
 GCATTGAATTACCGAACCGGTGTCCGCTTGCAGGGGGCCATCCTAACCATGGCATTAAAG  
 721 +-----+-----+-----+-----+-----+-----+ 780  
 CGTAACCTTAATGGCTTGGCCACAGGCGAACGCCCGGTAGGATTGGTACCGTAAATTC

a A L N Y R T G V R L R G A I L T M A F K .  
 AAGATCCTTAAGTAAAGAACATTAAAGAGAAATCCCTGGGTGAGCTCATCAACATTG  
 781 +-----+-----+-----+-----+-----+-----+ 840

**Figure 13B**

TTCTAGGAATTCAATTCTTGTAAATTCTCTTTAGGGACCCACTCGAGTAGTTGTAAACG

a K I L K L K N I K E K S L G E L I N I C .

TCCAACGATGGGCAGAGAATGTTGAGGCAGCAGCCGTTGGCAGCCTGCTGGCTGGAGGA  
841 -----+-----+-----+-----+-----+-----+ 900  
AGGTTGCTACCGTCTCTTACAAACTCCGTCGTCGGCAACCGTCGGACGACCGACCTCCT

a S N D G Q R M F E A A A V G S L L A G G .

CCCGTTGTTGCCATCTTAGGCATGATTATAATGTAATTATTCTGGGACCAACAGGCTTC  
901 -----+-----+-----+-----+-----+-----+ 960  
GGGCAACAACGGTAGAAATCCGTACTAAATATTACATTAATAAGACCCCTGGTTGCCGAAG

a P V V A I L G M I Y N V I I L G P T G F .

CTGGGATCAGCTGTTTATCCTCTTTACCCAGCAATGATGTTGCATCACGGCTCACA  
961 -----+-----+-----+-----+-----+-----+ 1020  
GACCCTAGTCGACAAAAATAGGAGAAAATGGGTGTTACTACAAACGTAGTGCCGAGTGT

a L G S A V F I L F Y P A M M F A S R L T .

GCATATTCAGGAGAAAATCGTGGCCGCCACGGATGAACGTGTCCAGAAGATGAATGAA  
1021 -----+-----+-----+-----+-----+ 1080  
CGTATAAAGTCCTCTTTACGCACCGGGCGGTGCCTACTTGACACAGGTCTTACTTACTT

a A Y F R R K C V A A T D E R V Q K M N E .

GTTCTTACTTACATTAATTTATCAAATGTATGCCCTGGTCAAAGCATTTCTCAGAGT  
1081 -----+-----+-----+-----+-----+ 1140  
CAAGAATGAATGTAATTAAATAGTTTACATACGGACCCAGTTGTAAAAGAGTCTCA

a V L T Y I K F I K M Y A W V K A F S O S .

GTTCAAGAAAATCCGCGAGGAGGGAGCGTCGGATATTGGAAAAAGCCGGTACTTCCAGGGT  
1141 -----+-----+-----+-----+-----+ 1200  
CAAGTCTTTAGGCCTCCCTCGCAGCCTATAACCTTTTGGCCCATGAAGGTCCCA

a V Q K I R E E E R R I L E K A G Y F Q G .

ATCACTGTGGGTGTGGCTCCCATGTGGTGGTATTGCCAGCGTGGTACCTCTGT  
1201 -----+-----+-----+-----+-----+ 1260  
TAGTGACACCCACACCGAGGGTAACACCACTAACGGTCGCACCACTGGAAGAGACAA

Figure 13C

27/56

a I T V G V A P I V V V I A S V V T F S V .

CATATGACCTGGGCTTCGATCTGACAGCAGCACAGGCTTCACAGTGGTACAGTCTTC  
 1261 -----+-----+-----+-----+-----+-----+-----+-----+ 1320  
 GTATACTGGGACCGAAGCTAGACTGTCGTCGTCCGAAAGTGTACCACTGTCAGAAG

a H M T L G F D L T A A Q A F T V V T V F .

AATCCATGACTTTGCTTGAAAGTAACACCGTTTCAGTAAAGTCCCTCTCAGAAGCC  
 1321 -----+-----+-----+-----+-----+-----+-----+ 1380  
 TTAAGGTACTGAAAACGAAACTTCATTGGCAAAAGTCATTCAGGGAGAGTCTTCGG

a N S M T F A L K V T P F S V K S L S E A .

TCAGTGGCTGTTGACAGATTAAGAGTTGTTCTAATGGAAGAGGTTACATGATAAAG  
 1381 -----+-----+-----+-----+-----+-----+-----+ 1440  
 AGTCACCGACAACGTCTAAATTCTCAAACAAAGATTACCTTCTCCAAGTGTACTATTC

a S V A V D R F K S L F L M E E V H M I K .

AACAAACCAGCCAGTCCTCACATCAAGATAGAGATGAAAAATGCCACCTGGCATGGAC  
 1441 -----+-----+-----+-----+-----+-----+-----+ 1500  
 TTGTTTGGTCGGTCAGGAGTGTAGTTCTATCTACTTTTACGGTGGAACCGTACCTG

a N K P A S P H I K I E M K N A T L A W D .

TCCTCCCACCTCCAGTATCCAGAACTGCCAACGCTGACCCCCAAATGAAAAAGACAAG  
 1501 -----+-----+-----+-----+-----+-----+-----+ 1560  
 AGGAGGGTGAGGTCTAGGTCTTGAGCGGGTTCGACTGGGGTTTACTTTCTGTTC

a S S H S S I Q N S P K L T P K M K K D K .

AGGGCTTCCAGGGCAAGAAAGAGAAGGTGAGGCAGCTGCAGCGCACTGAGCATCAGGCG  
 1561 -----+-----+-----+-----+-----+-----+-----+ 1620  
 TCCCGAAGGTCCCCGTTCTTCTCTTCCACTCCGTCACGTCGCGTACTCGTAGTCCGC

a R A S R G K K E K V R Q L Q R T E H Q A .

GTGCTGGCAGAGCAGAAAGGCCACCTCCTGGACAGTGACGAGCGGCCAGTCCGAA  
 1621 -----+-----+-----+-----+-----+-----+-----+ 1680  
 CACGACCGTCTCGTCTTCCGGTGGAGGAGGACCTGTCACTGCTCGCCGGTCAGGGCTT

Figure 13D

a V L A E Q K G H L L L D S D E R P S P E .  
 GAGGAAGAAGGCAAGCACATCCACCTGGGCCACCTGCGCTTACAGAGGACACTGCACAGC  
 1681 ----- + ----- + ----- + ----- + ----- + 1740  
 CTCCTTCTTCCGTTCTGTAGGTGGACCCGGTGGACGCGAATGTCTCCTGTGACGTGTCG  
  
 a E E E G K H I H L G H L R L O R T L H S .  
 ATCGATCTGGAGATCCAAGAGGGTAAACTGGTTGGAATCTGCGGCAGTGTGGGAAGTGGA  
 1741 ----- + ----- + ----- + ----- + ----- + 1800  
 TAGCTAGACCTCTAGGTTCTCCATTGACCAACCTTAGACGCCGTACACCCCTCACCT  
  
 a I D L E I Q E G K L V G I C G S V G S G .  
 AAAACCTCTCTCATTTCAGCCATTAGGCCAGATGACGCTTAGAGGGCAGCATTGCA  
 1801 ----- + ----- + ----- + ----- + ----- + 1860  
 TTTGGAGAGAGTAAAGTCGGTAAAATCCGGTCTACTGCGAAGATCTCCCGTGTAAACGT  
  
 K T S L I S A I L G Q M T L L E G S I A .  
 ATCA GTGGAACCTTCGTTATGTGGCCCAGCAGGCCTGGATCCTCAATGCTACTCTGAGA  
 1861 ----- + ----- + ----- + ----- + ----- + 1920  
 TAGTCACCTTGGAAAGCGAATAACCCGGTCTCCGGACCTAGGAGTTACGATGAGACTCT  
  
 I S G T F A Y V A Q Q A W I L N A T L R .  
 GACAACATCCTGTTGGGAAGGAATATGATGAAGAAAGATAACAACCTGTGCTGAACAGC  
 1921 ----- + ----- + ----- + ----- + ----- + 1980  
 CTGTTGAGGACAAACCCCTCCTATACTACTCTTCTATGTTGAGACACGACTTGTG  
  
 D N I L F G K E Y D E E R Y N S V L N S .  
 TGCTGCCTGAGGCCTGACCTGGCATTCTTCCCAGCAGCGACCTGACGGAGATTGGAGAG  
 1981 ----- + ----- + ----- + ----- + ----- + 2040  
 ACGACGGACTCCGGACTGGACCGGTAAGAAGGGTCGCTGGACTGCCTCTAACCTCTC  
  
 C C L R P D L A I L P S S D L T E I G E .  
 CGAGGAGCCAACCTGAGCGGTGGGCAGCGCCAGAGGATCAGCCTGCCCCGGCCTGTAT  
 2041 ----- + ----- + ----- + ----- + ----- + 2100  
 GCTCCTCGGTTGGACTGCCACCCGTCGCGGTCTCTAGTCGGAACGGGCCGGAACATA  
  
 R G A N L S G G O R Q R I S L A R A L Y .

**Figure 13E**

AGT GAC AGG AGG AGC AT CT AC AT C CT GG AC G AC C C C C T C AG T G C C T T A G AT G C C C AT G T G G G C  
 2101 ----- + ----- + ----- + ----- + ----- + ----- + 2160  
 T C A C T G T C C T C G T A G A T G T A G G A C C T G C T G G G G A G T C A C G G A A T C T A C G G G T A C A C C C G

**a S D R S I Y I L D D P L S A L D A H V G**

2161 AACCACATCTTCAATAGTGCTATCCGGAAACATCTCAAGTCCAAGACAGTTCTGTTGTT  
 2220 TTGGTGTAGAAGTTATCACGATAGGCCTTGTAGAGTTCAAGGTTCTGTCAAGACAAACAA

a N H I F N S A I R K H L K S K T V L E V

ACCCACCACTTACAGTACCTGGTTGACTGTGATGAAGTGATCTTCATGAAAGAGGGCTGT  
 2221 — + — + — + — + — + — + 2280  
 TGGGTGGTCAATGTCATGGACCAACTGACACTACTTCACTAGAAGTACTTCTCCGACA

**a** THQLQYLVDCDEVIEMKEGC

ATTACGGAAAGAGGCACCCATGAGGAACGTGATGAATTAAATGGTACTATGCTACCATT  
 2281 + + + + + + + + + 2340  
 TAATGCCTTCTCCGTGGTACTCCTTACTAAATTACCACTGATACGATGGTAA

a I T E R G T H E E L M N L N G D Y A T I -

2341 TTTAACCTGTTGGAGAGACACCGCCAGTTGAGATCAATTCAAAAAAGGAAACC  
 2400 AAATTATTGGACAACGACCCCTCTCTGTGGCGGTCAACTCTAGTTAAGTTTTTCCCTTGG

a F N N L L L G E T P P V E I N S K K E T .

2401 AGTGGTTCACAGAAGAAGTCACAAGACAAGGGTCCTAAAACAGGATCAGTAAAGAAGGAA  
 2460 TCACCAAGTGTCTTCTCAGTGTCTGTTCCAGGATTGTCCTAGTCATTCTTCCTT

a S G S Q K K S Q D K G P K T G S V K K E .

AAAGCAGTAAAGCCAGAGGAAGGGCAGCTTGTGCAGCTGGAAGAGAAAGGGCAGGGTTCA  
 2461 — + — + — + — + — + — + 2520  
 TTCGTCATTCGGTCTCCCTCCCGTCGAACACGTCGACCTTCTCTTCCCGTCCCAAGT

**a** K A V K P E E G Q L V Q L E E K G Q G S .

**Figure 13F**

GTGCCCTGGTCAGTATATGGTGTACATCCAGGCTGCTGGGGCCCTGGCATTCTG  
 2521 -----+-----+-----+-----+-----+-----+ 2580  
 CACGGGACCAAGTCATATACCAACAGATGTAGGTCCGACGACCCCCGGGAACCGTAAGGAC

a V P W S V Y G V Y I Q A A G G P L A F L .

GTTATTATGCCCTTTCATGCTGAATGTAGGCAGCACCGCCTCAGCACCTGGTGGTTG  
 2581 -----+-----+-----+-----+-----+-----+ 2640  
 CAATAATACCGGGAAAAGTACGACTTACATCGTCGTGGCGGAAGTCGTGGACCAAC

a V I M A L F M L N V G S T A F S T W W L .

AGTTACTGGATCAAGCAAGGAAGCGGGAACACCACTGTGACTCGAGGGAACGAGACCTCG  
 2641 -----+-----+-----+-----+-----+-----+ 2700  
 TCAATGACCTAGTCGTTCCCTCGCCCTGTGGTGACACTGAGCTCCCTGCTCTGGAGC

a S Y W I K Q G S G N T T V T R G N E T S .

GTGAGTGACAGCATGAAGGACAATCCTCATATGCACTATGCCAGCATCTACGCCCTC  
 2701 -----+-----+-----+-----+-----+-----+ 2760  
 CACTCACTGTCGTAACCTCGTTAGGAGTATACGTATGAGCTCGTAGATGCCAG

a V S D S M K D N P H M Q Y Y A S I Y A L .

TCCATGGCAGTCATGCTGATCCTGAAAGCCATTGAGGAGTTGTCTTGTCAAGGGCACG  
 2761 -----+-----+-----+-----+-----+-----+ 2820  
 AGGTACCGTCAGTACGACTAGGACTTCGGTAAGCTCCTCAACAGAAACAGTCCCGTGC

a S M A V M L I L K A I R G V V F V K G T .

CTGCGAGCTCCTCCGGCTGCATGACGAGCTTCCGAAGGATCCTCGAAGGCCCTATG  
 2821 -----+-----+-----+-----+-----+-----+ 2880  
 GACGCTCGAAGGAGGGCCGACGTACTGCTCGAAAAGGCTCCTAGGAAGCTCGGGATAC

a L R A S S R L H D E L F R R I L R S P M .

AAGTTTTTGACACGACCCCCACAGGGAGGATTCTCAACAGGTTTCAAAGACATGGAT  
 2881 -----+-----+-----+-----+-----+-----+ 2940  
 TTCAAAAAACTGTGCTGGGGGTGTCCCTCTAAGAGTTGTCCAAAAGGTTCTGTACCTA

a K F F D T T P T G R I L N R F S K D M D .

GAAGTTGACGTGCGGCTGCCGTTCCAGGCCGAGATGTTATCCAGAACGTTATCCTGGTG

Figure 13G

2941 ----- + ----- + ----- + ----- + ----- + 3000  
CTTCAACTGCACGCCGACGGCAAGGTCCGGCTCTACAAGTAGGTCTTGCAATAGGACCA

a E V D V R L P F Q A E M F I O N V I I V

TTCTTCTGTGTGGGAATGATCGCAGGAGTCTTCCCCTGGTTCTTGTGGCAGTGGGGCC  
 3001 ----- + ----- + ----- + ----- + ----- + 3060  
 AAGAAGACACACCCTTACTAGCGTCCTCAGAAGGGCACCAAGGAACACCCGACCCCCCCC

<sup>a</sup> F F C V G M I A G V E P W E I V A V C S

3061 CTTGTCATCCTCTTTCAGTCCTGCACATTGTCTCCAGGGTCTGATTGGGAGCTGAAG  
 3120 GAACAGTAGGAGAAAAGTCAGGACGTGAAACAGAGGTCCCAGGACTAAGCCCTCCACTTC

a L V I L F S V L H I V S B V I I R E I k

CGTCTGGACAATATCACGCAGTCACCTTCCTCTCCCACATCACGTCCAGCATAACAGGGC  
 3121 — + — + — + — + — + — + 3180  
 GCAGACCTGTTATAGTGCAGTGGAAAGGGAGAGGGGTGTAGTGCAGGTCGTATGTCCCC

a R L D N I T Q S P F L S H I T S S I O G

3181 CTTGCCACCATCCACGCCCTACAATAAAGGGCAGGAGTTCTGCACAGATACCAGGAGCTG  
 3240 GAACGGTGGTAGGTGCGGATGTTATTCGGCTCTCAAAGACGIGTCTATGGCTCTCCAC

a LATIHAYNKGOEEIHRYOEI

3241 — + — + — + — + — + 3300  
 GACCTACTGTTGGTTCGAGGAAAAAAAAACAAATGCACACGCTACGCCACCGACCGACAC

<sup>a</sup> L D D N Q A P F E I E T C A M B W I L A V

CGGCTGGACCTCATCAGCATGCCCTCATCACCAACCGGGGCTGATGATCGTTCTTATG  
 3301 — + — + — + — + — + 3360  
 GCCGACCTGGAGTAGTCGTAGCAGGAGTAGTGGTGGTGCCTGACTAAGCAAGAATAC

<sup>a</sup> R L D L I S I A L I T T T G I M I V I M

3361 CACGGGCAGATTCCCCAGCCTATGCGGGTCTGCCATCTCTTATGCTGTCCAGTTAACG 3420

Figure 13H

GTGCCGTCTAAGGGGCGGATACGCCAGAGCGGTAGAGAATACGACAGGTCAATTGC

a H G Q I P P A Y A G L A I S Y A V O L T

GGGCTGTTCCAGTTACGGTCAGACTGGCATCTGAGACAGAACGCTCGATTACCTCGGTG  
3421 -----+-----+-----+-----+-----+ 3480  
CCCGACAAGGTCAAATGCCAGTCTGACCGTAGACTCTGTCTCGAGCTAAGTGGAGCCAC

a G L F Q F T V R L A S E T E A R F T S V

GAGAGGATCAATCACTACATTAAGACTCTGCCTTGGAACGCACCTGCCAGAACATTAAGAAC  
3481 -----+-----+-----+-----+-----+ 3540  
CTCTCCTAGTTAGTGTATGTAATTCTGAGACAGGAACCTCGTGGACGGCTTAATTCTG

a E R I N H Y I K T L S L E A P A R I K N

AAGGCTCCCTCCCTGACTGGCCCCAGGAGGGAGAGGTGACCTTGAGAACGCAGAGATG  
3541 -----+-----+-----+-----+-----+ 3600  
TTCCGAGGGAGGGACTGACCGGGCTCCCTCCACTGGAAACTCTGCGTCTAC

a K A P S P D W P Q E G E V T F E N A E M

AGGTACCGAGAAAACCTCCCTTGTCTAAAGAAAGTATCCTCACGATCAAACCTAAA  
3601 -----+-----+-----+-----+-----+ 3660  
TCCATGGCTCTTGGAGGGAGAACAGGATTCTTCATAGGAAGTGCTAGTTGGATT

a R Y R E N L P L V L K K V S F T I K P K

GAGAAGATTGGCATTGTGGGGCGGACAGGATCAGGAAGTCCTCGCTGGGATGGCCCTC  
3661 -----+-----+-----+-----+-----+ 3720  
CTCTCTAACCGTAACACCCCCGCTGTCTAGTCCCTTCAGGAGCGACCCCTACCGGGAG

a E K I G I V G R T G S G K S S L G M A L

TTCCGTCTGGTGGAGTTATCTGGAGGCTGCATCAAGATTGATGGAGTGAGAACATCAGTGAT  
3721 -----+-----+-----+-----+-----+ 3780  
AAGGCAGACCACCTCAATAGACCTCCGACGTAGTTCTAACTACCTCACTCTAGTCACTA

a F R L V E L S G G C I K I D G V R I S D

ATTGGCCTTGGCACCTCCGAAGCAAACCTCTATCATTCCCTCAAGAGGCCGGTGTGTTC  
3781 -----+-----+-----+-----+-----+ 3840  
TAACCGGAACGGCTGGAGGCTTCGTTGAGAGATAGTAAGGAGTTCTGGCCACGACAAG

Figure 13I

a I G L A D L R S K L S I I P Q E P V L F .

AGTGGCACTGTCAGATCAAATTGGACCCCTCAACCAGTACACTGAAGACCAGATTTGG  
3841 -----+-----+-----+-----+-----+-----+ 3900  
TCACCGTGACAGTCTAGTTAACCTGGGAAGTGGTACGTGACTTCTGGCTAAACC

a S G T V R S N L D P F N O Y T E D O I W .

GATGCCCTGGAGAGGACACACATGAAAGAATGTATTGCTCAGCTACCTCTGAAACTTGAA  
3901 -----+-----+-----+-----+-----+-----+ 3960  
CTACGGGACCTCTCCTGTGTACTTTCTTACATAACGAGTCGATGGAGACTTGAACCT

a D A L E R T H M K E C I A Q L P L K L E .

TCTGAAGTGATGGAGAATGGGGATAACTTCTCAGTGGGGAACGGCAGCTTGTGCATA  
3961 -----+-----+-----+-----+-----+-----+ 4020  
AGACTTCACTACCTCTTACCCCTATTGAAGAGTCACCCCTTGCCGTCGAGAACACGTAT

a S E V M E N G D N F S V G E R Q L L C I .

GCTAGAGCCCTGCTCCGCCACTGTAAGATTCTGATTTAGATGAAGGCCACAGCTGCCATG  
4021 -----+-----+-----+-----+-----+-----+ 4080  
CGATCTCGGGACGAGGCGGTGACATTCTAAGACTAAAATCTACTTCGGTGTGACGGTAC

a A R A L L R H C K I L I L D E A T A A M .

GACACAGAGACAGACTTATTGATTCAAGAGACCATCCGAGAAGCATTGCAGACTGTACC  
4081 -----+-----+-----+-----+-----+-----+ 4140  
CTGTGTCTCTGTCTGAATAACTAAGTTCTGGTAGGCTCTCGTAAACGTCTGACATGG

a D T E T D L L I Q E T I R E A F A D C T .

ATGCTGACCATTGCCCATGCCCTGCACACGGTTAGGCTCCGATAGGATTATGGTGTG  
4141 -----+-----+-----+-----+-----+-----+ 4200  
TACGACTGGTAACGGGTAGCGGACGTGTGCCAAGATCCGAGGCTATCCTAATACCACGAC

a M L T I A H R L H T V L G S D R I M V L .

GCCCAGGGACAGGTGGAGTTGACACCCCATGGTCTGTCCAACGACAGTTCC  
4201 -----+-----+-----+-----+-----+-----+ 4260  
CGGGTCCCTGTCCACCAACCTCAAACGTGGGTAGCCAGGAAGACAGGTTGCTGTCAAGG

Figure 13J

34/56

a A Q G Q V V E F D T P S V L L S N D S S .

CGATTCTATGCCATGTTGCTGCTGCAGAGAACAAAGGTCGCTGTCAAGGGCTGA  
4261 ----- + ----- + ----- + ----- + ----- 4314  
GCTAAGATAACGGTACAAACGACGACGTCTTGTCCAGCGACAGTTCCGACT

a R F Y A M F A A A E N K V A V K G .

## Figure 13K

## MOAT D cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGGACGCCCTGTGGTCCGGGGAGCTCGGCTCCAAGTTCTGGACTCCAACCTGTCT  
 1 -----+-----+-----+-----+-----+ 60  
 TACCTGCGGGACACGCCAAGGCCCTCGAGCCGAGGTTAAGACCCCTGAGGTTGGACAGA  
  
 a M D A L C G S G E L G S K F W D S N L S .  
  
 GTGCACACAGAAAACCCGGACCTCACTCCCTGCTTCCAGAACTCCCTGCTGGCCTGGGTG  
 61 -----+-----+-----+-----+-----+ 120  
 CACGTGTGTCTTTGGGCCTGGAGTGAGGGACGAAGGTCTTGAGGGACGACCGGACCCAC  
  
 a V H T E N P D L T P C F Q N S L L A W V .  
  
 CCCTGCATCTACCTGTGGTCGCCCTGCCCTGCTACTTGCTCTACCTGCGGCACCATTGT  
 121 -----+-----+-----+-----+-----+ 180  
 GGGACGTAGATGGACACCCAGCGGGACGGACGATGAACGAGATGGACGCCGTGGTAACA  
  
 a P C I Y L W V A L P C Y L L Y L R H H C .  
  
 CGTGGCTACATCATCCTCTCCACCTGTCCAAGCTCAAGATGGCCTGGGTGTCTGCTG  
 181 -----+-----+-----+-----+-----+ 240  
 GCACCGATGTAGTAGGAGAGGGTGGACAGGTTCGAGTTCTACCAGGACCCACAGGACGAC  
  
 a R G Y I I L S H L S K L K M V L G V L L .  
  
 TGGTGCCTCCCTGGCGGACCTTTTACTCCTCATGGCCTGGTCCATGGCCGGGCC  
 241 -----+-----+-----+-----+-----+ 300  
 ACCACGCAGAGGACCCGCTGGAAAAAAATGAGGAAGGTACCGGACCAGGTACCGGCCGG  
  
 a W C V S W A D L F Y S F H G L V H G R A .  
  
 CCTGCCCTGTTTCTTGTCAACCCCTTGGTGGTGGGTCACCATGCTGGCCACC  
 301 -----+-----+-----+-----+-----+ 360  
 GGACGGGGACAAAAGAAACAGTGGGGAAACCACCAACCCAGTGGTACGACGACCGGTGG  
  
 a P A P V F F V T P L V V G V T M L L A T .  
  
 CTGCTGATAACAGTATGAGCGGCTGCAGGGCGTACAGTCTCGGGGTCTCATTATCTTC

Figure 14A

**Figure 14B**

TCCTTCGTCTTTCGTCTGCCGTGCTGTGTTCCGAAGTCGTGCGGACCCCTTTACGG

a R K Q E K Q T A R H K A S A A P G K N A -

TCCGGCGAGGACGAGGTGCTGCTGGGTGCCGCCAGGCCCCGAAGCCCTCCTCCTG  
841 ----- + ----- + ----- + ----- + ----- + 900  
AGGCCGCTCCTGCTCCACGACGACCCACGGGCCGGTCCGGGCCTCGGGAGGAAGGAC

a S G E D E V L L G A R P R P R K P S F L -

AAGGCCCTGCTGCCACCTCGGCTCCAGCTCCTCATCAGTGCCTGCTCAAGCTTATC  
901 ----- + ----- + ----- + ----- + ----- + 960  
TTCCGGGACGACCGGTGGAAGCCGAGGTGCAAGGAGTAGTCACGGACGAAGTCGAATAG

a K A L L A T F G S S F L I S A C F K L I -

CAGGACCTGCTCTCCTCATCAATCCACAGCTGCTCAGCATCCTGATCAGGTTATCTCC  
961 ----- + ----- + ----- + ----- + ----- + 1020  
GTCCTGGACGAGAGGAAGTAGTTAGGTGTCGACGAGTCGTAGGACTAGTCAAATAGAGG

a Q D L L S F I N P Q L L S I L I R F I S -

AACCCCATGGCCCCCTCCTGGGGCTCCTGGTGGCTGGCTGATGTTCTGTGCTCC  
1021 ----- + ----- + ----- + ----- + ----- + 1080  
TTGGGGTACCGGGGGAGGACCAACCCGAAGGACCAACCGACCCGACTACAAGGACACGAGG

a N P M A P S W W G F L V A G L M F L C S -

ATGATGCAGTCGCTGATCTAACACTATTACCACTACATCTTGACTGGGTGAAG  
1081 ----- + ----- + ----- + ----- + ----- + 1140  
TACTACGTCAAGCAGTAGAATGTTGATAATGGTATGTAGAACACTGACCCACTTC

a M M Q S L I L Q H Y Y H Y I F V T G V K -

TTTCGTACTGGGATCATGGGTGTCATCTACAGGAAGGCTCTGGTTATACCAACTCAGTC  
1141 ----- + ----- + ----- + ----- + ----- + 1200  
AAAGCATGACCTAGTACCCACAGTAGATGTCCTCCGAGACCAATAGTGGTTGAGTCAG

a F R T G I M G V I Y R K A L V I T N S V -

AAACGTGCGTCCACTGTGGGGAAATTGTCAACCTCATGTCAGTGGATGCCAGCGCTTC  
1201 ----- + ----- + ----- + ----- + ----- + 1260  
TTTGCACGCAGGTGACACCCCTTAACAGTTGGAGTACAGTCACCTACGGGTCGCGAAG

Figure 14C

a K R A S T V G E I V N L M S V D A Q R F .

ATGGACCTTGCCTCTCAATCTGCTGTGGTCAGCACCCCTGCAGATCATCCTGGCG  
1261 -----+-----+-----+-----+-----+ 1320  
TACCTGGAACGGGGAGGGAGTTAGACGACACCAGTCGTGGGACGTCTAGTAGGACCGC

a M D L A P F L N L L W S A P L Q I I L A .

ATCTACTTCCTCTGGCAGAACCTAGGTCCCTCTGCTCTGGCTGGAGTCGCTTCATGGTC  
1321 -----+-----+-----+-----+-----+ 1380  
TAGATGAAGGAGACCGCTGGATCCAGGGAGACAGGACCGACCTCAGCGAAAGTACCAAG

a I Y F L W Q N L G P S V L A G V A F M V .

TTGCTGATTCCACTCAACGGAGCTGTGGCCGTGAAGATGCGCGCCTCCAGGTAAAGCAA  
1381 -----+-----+-----+-----+-----+ 1440  
AACGACTAAGGTGAGTTGCCTCGACACCCGGCACTCTACGCGCGGAAGGTCCATTCGTT

a L L I P L N G A V A V K M R A F Q V K Q .

ATGAAATTGAAGGACTCGCGCATCAAGCTGATGAGTGAGATCCTGAACGGCATCAAGGTG  
1441 -----+-----+-----+-----+-----+ 1500  
TACTTTAACCTCCTGAGCGCGTAGTCGACTACTCACTCTAGGACTTGGCTAGTTCCAC

a M K L K D S R I K L M S E I L N G I K V .

CTGAAGCTGTACGCCTGGAGCCCAGCTTCCTGAAGCAGGTGGAGGGCATCCGGCAGGGT  
1501 -----+-----+-----+-----+-----+ 1560  
GACTTCGACATGCGGACCTCGGGTCAAGGACTTCGTCCACCTCCCGTAGGCCGTCCCA

a L K L Y A W E P S F L K Q V E G I R Q G .

GAGCTCCAGCTGCTGCGCACGGCGGCCTACCTCCACACCAACACCACCTTCACCTGGATG  
1561 -----+-----+-----+-----+-----+ 1620  
CTCGAGGTGACGACCGTGCCTGGATGGAGGTGGTGGTGGAGTGGACCTAC

a E L Q L L R T A A Y L H T T T F T W M .

TGCAGCCCCCTTCTGGTACCCCTGATCACCCCTGGGTGTACGTGTACGTGGACCCAAAC  
1621 -----+-----+-----+-----+-----+ 1680  
ACGTCGGGGAGGACCACTGGGACTAGTGGAGACCCACATGCACATGCACCTGGTTG

Figure 14D

a C S P F L V T L I T L W V Y V Y V D P N -

1681 AATGTGCTGGACGCCGAGAAGGCCTTGTCTGTCTGCTTAAATATCTTAAGACTT  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
TTACACGACCTGCGCTCTCCGGAAACACAGACACAGGAACAAATTATAGAATTCTGAA

a N V L D A E K A F V S V S L F N I L R L -

1741 CCCCTAACATGCTGCCCGAGTTAACAGAACCTGACTCAGGCCAGTGTGTCTGAAA  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800  
GGGGAGTTGTACGACGGGTCAATTAGTCGTTGGACTGAGTCCGGTCACACAGAGACTT

a P L N M L P Q L I S N L T Q A S V S L K -

1801 CGGATCCAGCAATTCTGAGCCAAGAGGAACCTGACCCCCAGAGTGTGGAAAGAAAGACC  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860  
GCCTAGGTGTTAAGGACTCGTTCTCCTGAACCTGGGGTCTCACACCTTCTTCTGG

a R I Q Q F L S Q E E L D P Q S V E R K T -

1861 ATCTCCCCAGGCTATGCCATCACCATACACAGTGGCACCTCACCTGGGCCAGGACCTG  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920  
TAGAGGGGTCCGATACTGGTAGTGGTATGTGTACCGTGGAAAGTGGACCCGGGTCTGGAC

a I S P G Y A I T I H S G T F T W A Q D L -

1921 CCCCCCACTCTGCACAGCCTAGACATCCAGGTCCCGAAAGGGGACTGGTGGCCGTGGTG  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
GGGGGGTGAGACGTGTCGGATCTGTAGGTCCAGGGCTTCCCGTGACCACCGGACACAC

a P P T L H S L D I Q V P K G A L V A V V -

1981 GGGCCTGTGGCTGTGGGAAGTCCTCCCTGGTGTCTGCCCTGCTGGAGAGATGGAGAAG  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040  
CCCGGACACCCGACACCCCTCAGGAGGGACCACAGACGGGACGACCCCTCTACCTCTTC

a G P V G C G K S S L V S A L L G E M E K -

2041 CTAGAAGGCAAAGTGCACATGAAGGCATGGATCCAGAACTGCACCTTCAGGAAACGTG  
+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
GATCTTCCGTTCACGTGACTTCCGTACCTAGGTCTTGACGTGAGAAGTCCTTGCAC

a L E G K V H M K A W I Q N C T L Q E N V -

## Figure 14E

SUBSTITUTE SHEET (RULE 26)

CTTTCGGCAAAGCCCTGAACCCCAAGCGCTACCAGCAGACTCTGGAGGCCTGTGCCTTG  
 2101 -----+-----+-----+-----+-----+-----+ 2160  
 GAAAAGCCGTTCGGGACTTGGGGTCGCATGGTCGCTGAGACCTCCGGACACGGAAC

a L F G K A L N P K R Y Q Q T L E A C A L .

CTAGCTGACCTGGAGATGCTGCCTGGTGGGATCAGACAGAGATTGGAGAGAAGGGCATT  
 2161 -----+-----+-----+-----+-----+-----+ 2220  
 GATCGACTGGACCTCTACGACGGACCACCCCTAGTCTGTCTAACCTCTTCCCGTAA

a L A D L E M L P G G D O T E I G E K G I .

AACCTGTCTGGGGGCCAGCGGCAGCGGTCACTGGCTCGAGCTGTTACAGTGATGCC  
 2221 -----+-----+-----+-----+-----+-----+ 2280  
 TTGGACAGACCCCCGGTCGCCGCCCCAGTCAGACCGAGCTGACAAATGTCACTACGG

a N L S G G Q R Q R V S L A R A V Y S D A .

GATATTTCTTGGATGACCCACTGTCCGCGGTGGACTCTCATGTGCCAAGCACATC  
 2281 -----+-----+-----+-----+-----+-----+ 2340  
 CTATAAAAGAACGACCTACTGGGTGACAGGCGCCACCTGAGAGTACACCGGTTGTAG

a D I F L L D D P L S A V D S H V A K H I .

TTTGACCACGTACGGCCAGAAGGCAGCTGGCAGGCAAGACCGAGTGCTGGTACG  
 2341 -----+-----+-----+-----+-----+-----+ 2400  
 AAACCTGGTGCAGTAGCCCGGTCTCCGCACGACCGTCCGTTCTGCCTCACGACCACTGC

a F D H V I G P E G V L A G K T R V L V T .

CACGGCATTAGCTTCCCTGCCCGAGACAGACTTCATCATTGTGCTAGCTGATGGACAGGTG  
 2401 -----+-----+-----+-----+-----+-----+ 2460  
 GTGCCGTAATCGAAGGACGGGTCTGTCTGAAGTAGTAACACGATCGACTACCTGTCCAC

a H G I S F L P Q T D F I I V L A D G Q V .

TCTGAGATGGGCCGTACCCAGCCCTGCTGCAGCGCAACGGCTCTTGCCAACCTTCTC  
 2461 -----+-----+-----+-----+-----+-----+ 2520  
 AGACTCTACCCGGGCATGGTCGGGACGACGTGCGTGGCAGGAAACGGTTGAAAGAG

a S E M G P Y P A L L Q R N G S F A N F L .

Figure 14F

TGCAACTATGCCCGATGAGGACCAAGGGCACCTGGAGGAACAGCTGGACCGCGTTGGAA  
 2521 -----+-----+-----+-----+-----+-----+ 2580  
 ACGTTGATACGGGGCTACTCCTGGTCCCGTGGACCTCCTGTCGACCTGGCGAACCTT

a C N Y A P D E D Q G H L E D S W T A L E .

GGTGCAGAGGATAAGGAGGCAGTGCTGATTGAAGACACACTCAGCAACCACACGGATCTG  
 2581 -----+-----+-----+-----+-----+-----+ 2640  
 CCACGTCTCCTATTCCCTCCGTGACGACTAACTTCTGTGTGAGTCGTTGGTGTGCCTAGAC

a G A E D K E A L L I E D T L S N H T D L .

ACAGACAATGATCCAGTCACCTATGTGGTCCAGAAGCAGTTATGAGACAGCTGAGTGCC  
 2641 -----+-----+-----+-----+-----+-----+ 2700  
 TGTCTGTTACTAGGTCACTGGATACACCAGGTCTCGTCAAATACTCTGTCGACTCACGG

a T D N D P V T Y V V Q K Q F M R Q L S A .

CTGTCCCTCAGATGGGAGGGACAGGGCGGCTGTACCCCGGAGGCACCTGGTCCATCA  
 2701 -----+-----+-----+-----+-----+-----+ 2760  
 GACAGGAGTCTACCCCTCCCTGTCCCAGCCGGACATGGGGCCTCCGTGGACCCAGGTAGT

a L S S D G E G Q G R P V P R R H L G P S .

GAGAAGGTGCAGGTGACAGAGGCGAAGGCAGATGGGGCACTGACCCAGGAGGAGAAAGCA  
 2761 -----+-----+-----+-----+-----+-----+ 2820  
 CTCTTCCACGTCCACTGTCTCCGCTTCCGTCTACCCGTGACTGGTCCCTCTTCGT

a E K V Q V T E A K A D G A L T Q E E K A .

GCCATTGGCACTGTGGAGCTCAGTGTCTGGATTATGCCAAGGCCGTGGGCTCTGT  
 2821 -----+-----+-----+-----+-----+-----+ 2880  
 CGGTAACCGTGACACCTCGAGTCACACAAGACCCATAATACGGTCCGGCACCCGAGACA

a A I G T V E L S V F W D Y A K A V G L C .

ACCACGCTGCCATCTGTCTCTGTATGTGGTCAAAGTGCCTGCCATTGGAGCCAAT  
 2881 -----+-----+-----+-----+-----+-----+ 2940  
 TGGTGCACCGTAGACAGAGGACATACACCCAGTTACGCCGACGGTAACCTCGGTTA

a T T L A I C L L Y V G Q S A A A I G A N

GTGTGGCTCAGTGCCTGGACAAATGATGCCATGGCAGACAGTAGACAGAACACACTTCC

## Figure 14G

2941 ----- + ----- + ----- + ----- + ----- + 3000  
CACACCGAGTCACGGACCTGTTACTACGGTACCGTCTGTCATCTGTCCTTGTTGTGAAGG

a V W L S A W T N D A M A D S R O N N T S .

CTGAGGCTGGCGTCTATGCTGCTTAGGAATTCTGCAAGGGTCTTGGTATGCTGGC  
3001 -----+-----+-----+-----+-----+-----+ 3060  
GAATCCGACCCGAGATAACGACGAAATCCTAAGACGTTCCCCAAGAACCAACTACGGACCC

<sup>a</sup> L R L G V Y A A L G I I O G E I V M I A

<sup>a</sup> A M A M A A G G I O A A B V I H O A I I

a H N K I B S P O S F E D T T P C C C A

AACTGTTCTCCAAGGACATCTATGTCGTTGAGGTTCTGGCCCTGTCATCCTCATG  
 3181 — + — + — + — + — + 3240  
 TTGACGAAGAGGTCTCTGTAGATACAGCAACTCTGAAAGACCCCCCCAGACTTACGACTA

<sup>a</sup> N C E S K P I X V V D E V I A S V A M M

## **L I N S E E N A I S T I V V I M A S T E**

CTCTTCACTGTGGTCATCCTGCCCTGGCTGTGCTCTACACCTTAGTGCAGCGCTTCTAT  
 3301 — + — + — + — + — + 3360  
 GAGAAGTGACACCACTAGGACGGGGACCCACACGATCTGAAATGCGCTGGGAGCTT

<sup>a</sup> LET V VIII B1 A V L Y T L V C B S V

GCAGGCCACATCACGGCAACTGAAGCGGCTGGAATCAGTCAGCCGCTCACCTATCTACTCC  
3361 — + — + — + — + — + 3420

**Figure 14H**

CGTCGGTAGTGCCGTTGACTTCGCCGACCTAGTCAGTCGGCGAGTGGATAGATGAGG

a A A T S R Q L K R L E S V S R S P I Y S .

CACTTTCGGAGACAGTGACTGGTGCCAGTGTCACTCGGGCCTACAACCGCAGCCGGAT  
3421 ..... + ..... + ..... + ..... + ..... + 3480  
GTGAAAAGCCTCTGTCACTGACCACGGTCACAGTAGGCCGGATGTTGGCGTCGGCCCTA

a H F S E T V T G A S V I R A Y N R S R D .

TTTGAGATCATCAGTGATACTAAGGTGGATGCCAACCAGAGAAGCTGCTACCCCTACATC  
3481 ..... + ..... + ..... + ..... + ..... + 3540  
AAACTCTAGTAGTCACTATGATTCCACCTACGGTTGGTCTCTCGACGATGGGATGTAG

a F E I I S D T K V D A N Q R S C Y P Y I .

ATCTCCAACCAGGTGGCTGAGCATCGGAGTGGAGTTCTGGGGAACTGCGTGGTGCCTTT  
3541 ..... + ..... + ..... + ..... + ..... + 3600  
TAGAGGTTGGCCACCGACTCGTAGCCTCACCTCAAGCACCCCTTGACGCACCACGAGAAA

a I S N R W L S I G V E F V G N C V V L F .

GCTGCACTATTGCCGTATCGGGAGGAGCAGCCTGAACCCGGGGCTGGTGGCCTTCT  
3601 ..... + ..... + ..... + ..... + ..... + 3660  
CGACGTGATAAACGGCAGTAGCCCTCTCGTGGACTTGGCCCCGACCACCGGAAAGA

a A A L F A V I G R S S L N P G L V G L S .

GTGTCCACTCCTTGCAGGTGACATTGCTCTGAACCTGGATGATCGAATGATGTCAGAT  
3661 ..... + ..... + ..... + ..... + ..... + 3720  
CACAGGATGAGGAACGTCCACTGTAAACGAGACTTGACCTACTATGCTTACTACAGTCTA

a V S Y S L Q V T F A L N W M I R M M S D .

TTGGAATCTAACATCGTGGCTGTGGAGAGGGTCAAGGAGTACTCCAAGACAGAGACAGAG  
3721 ..... + ..... + ..... + ..... + ..... + 3780  
AACCTTAGATTGTAGCACCGACACCTCTCCAGTTCTCATGAGGTTCTGTCTGTCTC

a L E S N I V A V E R V K E Y S K T E T E .

GCGCCCTGGGTGGAGGCAAGCCGCCCTCCGAAGGTTGGCCCCACGTGGGGAGGTG  
3781 ..... + ..... + ..... + ..... + ..... + 3840  
CGCGGGACCCACCCACCTCCGTCGGCGGGAGGGCTTCCAACCGGGGTGCACCCCTCCAC

Figure 14I

a A P W V V E G S R P P E G W P P R G E V .

GAGTTCCGGAATTATTCTGCGCTACCGGCCGGCCTAGACCTGGTCTGAGAGACCTG  
 3841 -----+-----+-----+-----+-----+-----+ 3900  
 CTCAGGCCTTAATAAGACACCGCGATGCCGGCCGGATCTGGACCACGACTCTGGAC

a E F R N Y S V R Y R P G L D L V L R D L .

AGTCTGCATGTCCACGGTGGCGAGAAGGTGGGATCGTGGGCCGCACTGGGCTGGCAAG  
 3901 -----+-----+-----+-----+-----+-----+ 3960  
 TCAGACGTACACGTGCCACCGCTTCCACCCCTAGCACCCGGCGTACCCGACCGTTC

a S L H V H G G E K V G I V G R T G A G K .

TCTTCCATGACCCTTGCGCTGTTCCGCATCCTGGAGGCGCAAAGGGTGAATCCGCATT  
 3961 -----+-----+-----+-----+-----+-----+ 4020  
 AGAAGGTACTGGAAACGGACAAGGCGTAGGACCTCCGCCGTTCCACTTAGGCGTAA

a S S M T L C L F R I L E A A K G E I R I .

GATGGCCTCAATGTGGCAGACATCGGCCTCCATGACCTGCGCTCTCAGCTGACCATCATC  
 4021 -----+-----+-----+-----+-----+-----+ 4080  
 CTACCGGAGTTACACCGTCTGTAGCCGGAGGTACTGGACGCGAGAGTCGACTGGTAGTAG

a D G L N V A D I G L H D L R S Q L T I I .

CCGCAGGACCCATCCTGTTCTGGGGACCCCTGCGCATGAACCTGGACCCCTCGGCAGC  
 4081 -----+-----+-----+-----+-----+-----+ 4140  
 GGCGTCTGGGTAGGACAAGAGCCCTGGACGCGTACTGGACCTGGGAAGCCGTG

a P Q D P I L F S G T L R M N L D P F G S .

TACTCAGAGGAGGACATTGGTGGCTTGGAGCTGTCCACCTGCACACGTTGTGAGC  
 4141 -----+-----+-----+-----+-----+-----+ 4200  
 ATGAGTCTCCTCTGTAAACCAACCCGAAACCTCGACAGGGTGGACGTGTGCAAACACTCG

a Y S E E D I W W A L E L S H L H T F V S .

TCCCAGCCGGCAGGCCTGGACTTCCAGTGCTCAGAGGGCGGGAGAATCTCAGCGTGGC  
 4201 -----+-----+-----+-----+-----+-----+ 4260  
 AGGGTCGGCCGTCCGGACCTGAAGGTACAGAGTCTCCGCCCTTTAGAGTCGACCCG

Figure 14J

a S Q P A G L D F Q C S E G G E N L S V G -

CAGAGGCAGCTCGTGTGCCTGGCCCGAGCCCTGCTCCGCAAGAGCCGCATCCTGGTTTTA  
 4261 -----+-----+-----+-----+-----+-----+ 4320  
 GTCTCCGTCGAGCACACGGACCGGGCTCGGGACGAGGCCTCTGGCGTAGGACCAAAAT

a Q R Q L V C L A R A L L R K S R I L V L -

GACGAGGCCACACCTGCCATCGACCTGGAGACTGACAACCTCATCCAGGCTACCATCCGC  
 4321 -----+-----+-----+-----+-----+-----+ 4380  
 CTGCTCCGGTGTGACGGTAGCTGGACCTCTGACTGTTGGAGTAGGTCCATGGTAGGCG

a D E A T A A I D L E T D N L I Q A T I R -

ACCCAGTTGATAACCTGCACTGTCCCTGACCATCGCACACCGGCTAACACTATCATGGAC  
 4381 -----+-----+-----+-----+-----+-----+ 4440  
 TGGGTCAAACATGGACGTGACAGGACTGGTAGCGTGTGGCCGAATTGTGATAGTACCTG

a T Q F D T C T V L T I A H R L N T I M D -

TACACCAGGGCCTGGCCTGGACAAAGGAGTAGTAGCTGAATTGATTCTCCAGCCAAC  
 4441 -----+-----+-----+-----+-----+-----+ 4500  
 ATGTGGTCCCAGGACCAGGACCTGTTCCCTCATCATCGACTAAACTAAGAGGTCGGTTG

a Y T R V L V L D K G V V A E F D S P A N -

CTCATTGCAGCTAGAGGCATCTTCTACGGATGCCAGAGATGCTGGACTTGCTAA  
 4501 -----+-----+-----+-----+-----+-----+ 4557  
 GAGTAACGTGATCTCCGTAGAAGATGCCCTACCGGTCTACGACCTGAACGGATT

a L I A A R G I F Y G M A R D A G L A -

Figure 14K

## MOAT E cDNA AND AMINO ACID SEQUENCE ENCODED THEREBY

ATGGCCGCGCCTGCTGAGCCCTGCGCGGGGAGGGGGTCTGGAACCAGACAGAGCCTGAA  
 1 -----+-----+-----+-----+-----+-----+ 60  
 TACCGGCGCGGACGACTCGGGACGCGCCCCGCCCCAGACCTGGTCTGTCGGACTT

a M A A P A E P C A G Q G V W N O T E P E -

CCTGCCGCCACCAGCCTGCTGAGCCTGTGCTTCTGAGAACAGCAGGGGTCTGGTACCC  
 61 -----+-----+-----+-----+-----+-----+ 120  
 GGACGGCGGTGGTGGACGACTCGGACACGAAGGACTCTTGTGTCGCCCCAGACCCATGGG

a P A A T S L L S L C F L R T A G V W V P -

CCCATGTACCTCTGGTCCTGGTCCATCTACCTCCTTCCATCCACCAACCATGGCCGG  
 121 -----+-----+-----+-----+-----+-----+ 180  
 GGGTACATGGAGACCCAGGAACCAGGGTAGATGGAGGAGAAGTAGGTGGTGGTACCGGCC

a P M Y L W V L G P I Y L L F I H H H G R -

GGCTACCTCCGGATGTCCCCACTCTCAAAGCCAAGATGGTGCTGGATTGCCCTCATA  
 181 -----+-----+-----+-----+-----+-----+ 240  
 CCGATGGAGGCCTACAGGGTAGAAAGTTGGTTCTACCACGAACCTAACGGAGTAT

a G Y L R M S P L F K A K M V L G F A L I -

GTCCTGTGTACCTCCAGCGTGGCTGCGCTTTGGAAAATCCAACAGGAACGCCGTGAG  
 241 -----+-----+-----+-----+-----+-----+ 300  
 CAGGACACATGGAGGTGCGACCGACAGCGAGAACCTTTAGGTTGCCCTGGACTC

a V L C T S S V A A V A L W K I Q Q G T P E -

GCCCCAGAATTCTCATTCTACTGTGTGGCTACCGATGAGCTTCGAGTGTTC  
 301 -----+-----+-----+-----+-----+-----+ 360  
 CGGGGTCTTAAGGAGTAAGTAGGATGACACACCGAGTGGTGCTACTCGAACGCTAACAG

a A P E F L I H P T V W L T T M S F A V F -

CTGATTACACCCGAGAGGAAAAAGGGAGTCCAGTCATCTGGAGTGCTGGTTACTGG  
 361 -----+-----+-----+-----+-----+-----+ 420  
 GACTAAGTGTGGCTCTCTTCCCTCAGGTCAAGTCAGTAGACCTCACGACAAACCAATGACC

Figure 15A

a L I H T E R K K G V Q S S G V L F G Y W .  
 CTTCTCTGCTTGCTTGCAGCTACCAACGCTGCCAGCAGGCCTCCGGAGCGGGCTTC  
 421 -----+-----+-----+-----+-----+-----+ 480  
 GAAGAGACGAAACAGAACGGTCGATGGTGCACGGTCGTCCGGAGGCCTGCCCGAAG

a L L C F V L P A T N A A Q Q A S G A G F .  
 CAGAGCGACCCCTGTCCGCCACCTGCCACCTACCTATGCCGTCTGGTGGTGGCACAG  
 481 -----+-----+-----+-----+-----+-----+ 540  
 GTCTCGCTGGGACAGGCAGGTGGACAGGTGGATGGATACGGACAGAGACCACCGTGTGTC

a Q S D P V R H L S T Y L C L S L V V A Q .  
 TTTGTGCTGTCCTGCCCTGGCGGATCAACCCCCCTTCTCCCTGAAGACCCCCAGCAGTCT  
 541 -----+-----+-----+-----+-----+-----+ 600  
 AACACAGCACAGGACGGACCCGCTAGTTGGGGGAAGAAGGGACTCTGGGGTCGTCAAGA

a F V L S C L A D Q P P F F P E D P Q Q S .  
 AACCCCTGTCCAGAGACTGGGGCAGCCTCCCTCAAAGCCACGTTCTGGTGGTTCT  
 601 -----+-----+-----+-----+-----+-----+ 660  
 TTGGGGACAGGTCTCTGACCCCGTCGAAGGGAGGTTGGTGCAAGACCACCAAAGA

a N P C P E T G A A F P S K A T F W W V S .  
 GGCCTGGTCTGGAGGGATAACAGGAGGCCACTGAGACCAAAAGACCTCTGGTCGCTTGGG  
 661 -----+-----+-----+-----+-----+-----+ 720  
 CCGGACCAAGACCTCCCTATGCCCTGGTACTCTGGTTCTGGAGACCAAGCGAACCC

a G L V W R G Y R R P L R P K D L W S L G .  
 AGAGAAAACCTCTCAGAAGAACCTGTTCCGGCTTGAAAAGGAGTGGATGAGGAACCGC  
 721 -----+-----+-----+-----+-----+-----+ 780  
 TCTCTTTGAGGAGTCTCTGAACAAAGGGCGAACCTTCTCACCTACTCCTGGCG

a R E N S S E E L V S R L E K E W M R N R .  
 AGTGCAGCCGGAGGCACAACAAGGCAATAGCATTAAAAGGAAAGGCGGCAGTGGCATG  
 781 -----+-----+-----+-----+-----+-----+ 840  
 TCACGTGGGCCCTCCGTGTTCCGTTATCGTAAATTCTCCTCCGCCGTACCGTAC

Figure 15B

a S A A R R H N K A I A F K R K G G S G M .

AAGGCTCCAGAGACCGAGCCCTTCTACGGCAAGAAGGGAGCCAGTGGCGCCACTGCTG  
841 -----+-----+-----+-----+-----+-----+ 900  
TTCCGAGGTCTCTGGCTCGGAAGGATGCCGTTCTCCCTCGGTACCGCGGGTGACGAC

a K A P E T E P F L R Q E G S O W R P L L .

AAGGCCATCTGGCAGGTGTTCCATTCTACCTCCTCCTGGGGACCTCAGCCTCATCATC  
901 -----+-----+-----+-----+-----+-----+ 960  
TTCCGGTAGACCGTCCACAAGGTAAGATGGAAGGAGGACCCCTGGAGTCGGAGTAGTAG

a K A I W Q V F H S T F L L G T L S L I I .

AGTGATGTCTTCAGGTTCACTGTCCCCAAGCTGCTCAGCCTTCTGGAGTTATTGGT  
961 -----+-----+-----+-----+-----+-----+ 1020  
TCACTACAGAAGTCCAAGTGACAGGGGTCGACGAGTCGGAAAAGGACCTCAAATAACCA

a S D V F R F T V P K L L S L F L E F I G .

GATCCAAGCCTCCAGCCTGGAAAGGGTACCTCCTCGCCGTGCTGATGTTCTCTCAGCC  
1021 -----+-----+-----+-----+-----+-----+ 1080  
CTAGGGTTGGAGGTCGGACCTTCCGATGGAGGAGCGGCACGACTACAAGGAGAGTCGG

a D P K P P A W K G Y L L A V L M F L S A .

TGCCCTGCAAACGCTGTTGAGCAGCAGAACATGTACAGGCTAAGGTGCCGCAGATGAGG  
1081 -----+-----+-----+-----+-----+-----+ 1140  
ACGGACGTTGCGACAAACTCGTCGTCTGTACATGTCCGAGTCCACGGCGTCACTCC

a C L Q T L F E Q Q N M Y R L K V P Q M R .

TTGCGGTGCCATCACTGGCCTGGTGTACAGAAAGGTCTGGCTCTGTCCAGCGGCTCC  
1141 -----+-----+-----+-----+-----+-----+ 1200  
AACGCCAGCCGGTAGTGACCGGACCACATGTCTTCCAGGACCGAGACAGGTCGCCGAGG

a L R S A I T G L V Y R K V L A L S S G S .

AGAAAGGCCAGTGCAGGTGGGTGATGTGGTCAATCTGGTGTCCGTGGACGTGCAGCGGCTG  
1201 -----+-----+-----+-----+-----+-----+ 1260  
TCTTCCGGTCACGCCACCCACTACACCAGTTAGACCACAGGCACCTGCACGTCGCCGAC

R K A S A V G D V V N L V S V D V O R I .

Figure 15C

ACCGAGAGCGTCCTCACCTAACGGGCTGTGGCTGCCTCTCGTCTGGATCGTGGCTGC  
 1261 -----+-----+-----+-----+-----+-----+ 1320  
 TGGCTCTCGCAGGAGATGGAGTTGCCGACACCGACGGAGAGCAGACCTAGCACCAGACG

a T E S V L Y L N G L W L P L V W I V V C .

TTCGTCTATCTCTGGCAGCTCCTGGGGCCCTCCGCCCTCACTGCCATCGCTGTCTTCTG  
 1321 -----+-----+-----+-----+-----+-----+ 1380  
 AAGCAGATAGAGACCGTCGAGGACCCGGAGGCAGGAGTACGGTAGCGACAGAAGGAC

a F V Y L W Q L L G P S A L T A I A V F L .

AGCCTCCTCCCTCTGAATTCTTCATCTCCAAGAAAAGGAACCACCATCAGGAGGAGCAA  
 1381 -----+-----+-----+-----+-----+-----+ 1440  
 TCGGAGGGAGGGAGACTAAAGAAGTAGAGGTTCTTCTGGTAGTCCCTCGTT

a S L L P L N F F I S K K R N H H Q E E Q .

ATGAGGCAGAAGGACTCACGGGCACGGCTACCAAGCTATCCTCAGGAACCTCGAACAGACC  
 1441 -----+-----+-----+-----+-----+-----+ 1500  
 TACTCCGTCTCCTGAGTGCCCGTGGTCGAGATAGGAGTCCTTGAGCTTCTGG

a M R Q K D S R A R L T S S I L R N S K T .

ATCAAGTTCCATGGCTGGGAGGGAGCCTTCTGGACAGAGTCCTGGCATCCGAGGCCAG  
 1501 -----+-----+-----+-----+-----+-----+ 1560  
 TAGTTCAAGGTACCGACCCCTCCCTCGAAAGACCTGTCTCAGGACCCGTAGGCTCCGGTC

a I K F H G W E G A F L D R V L G I R G Q .

GAGCTGGCGCCTTGGACCTCCGGCTCCTCTGTGTGCGCTGGTCTTCAA  
 1561 -----+-----+-----+-----+-----+-----+ 1620  
 CTCGACCCCGGAAACGCCTGGAGGCCGGAGGAGAAGAGACACAGCGACCACAGGAAGGTT

a E L G A L R T S G L L F S V S L V S F Q .

GTGTCTACATTCTGGTCGCACTGGTGGTGGCTGTCCACACTCTGGTGGCCGAGAAT  
 1621 -----+-----+-----+-----+-----+-----+ 1680  
 CACAGATGTAAAGACCAGCGTGACCACACAAACGACAGGTGTGAGACCACCGGCTCTTA

a V S T F L V A L V V F A V H T L V A E N .

**Figure 15D**

1681 GCTATGAATGCAGAGAAAGCCTTGTGACTCTCACAGTTCTAACATCCTAACACAAGGCC  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
 CGATACTTACGTCTTTGGAAACACTGAGAGTGTCAAGAGTTGAGGAGTTGTTCCGG

a A M N A E K A F V T L T V L N I L N K A .

1741 CAGGCTTCTGCCCTTCTCCATCCACTCCCTCGTCCAGGCCGGTGTCTTGACCGT  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800  
 GTCCGAAAGGACGGGAAGAGGTAGGTGAGGGAGCAGGTCCGGGCCACAGGAAACTGGCA

a Q A F L P F S I H S L V Q A R V S F D R .

1801 CTGGTCACCTTCTGCCTGGAAAGAAGTTGACCCCTGGTGTCTAGACTCAAGTTCT  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860  
 GACCAGTGGAAAGGAGACGGACCTCTCAACTGGGACCACAGCATCTGAGTTCAAGGAGA

a L V T F L C L E E V D P G V V D S S S .

1861 GGAAGCGCTGCCGGAAAGGATTGCATCACCATACACAGTGCACCTCGCCTGGTCCCAG  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920  
 CCTTCGCGACGGCCCTTCTAACGTAGTGGTATGTGTACCGTGGAAAGCGGACCAGGGTC

a G S A A G K D C I T I H S A T F A W S Q .

1921 GAAAGCCCTCCCTGCCTCACAGAATAAACCTCACGGTCCCCAGGGCTGTCTGGCT  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
 CTTTCGGGAGGGACGGAGGTGTCTTATGGAGTGCCACGGGTCCGACAGACGACCGA

a E S P P C L H R I N L T V P Q G C L L A .

1981 GTTGTGGTCCAGTGGGGCAGGGAAAGTCCTCCCTGCTGTCCGCCCTCTGGGAGCTG  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040  
 CAACAGCCAGGTACCCCCGTCCCTCAGGAGGGACGACAGGCGGGAGGAACCCCTCGAC

a V V G P V G A G K S S L L S A L L G E L .

2041 TCAAAGGTGGAGGGGTTCGTGAGCATCGAGGGTGTGGCCTACGTGCCAGGAGGCC  
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
 AGTTTCCACCTCCCCAACGCACTCGTAGCTCCACGACACCGGATGCACGGGTCCCTCGG

a S K V E G F V S I E G A V A Y V P Q E A .

TGGGTGCAGAACACCTCTGTGGTAGAGAATGTGTGCTTCGGCAGGAGCTGGACCCACCC

**Figure 15E**

2101 ----- + ----- + ----- + ----- + ----- + 2160  
 ACCCACGTCTTGTGGAGACACCATCTTACACACAGAAGCCGTCCTGACCTGGGTGGG

a W V Q N T S V V E N V C F G Q E L D P P .

TGGCTGGAGAGAGTACTAGAAGCCTGTGCCCTGCAGCCAGATGTGGACAGCTCCCTGAG  
 2161 ----- + ----- + ----- + ----- + ----- + 2220  
 ACCGACCTCTCATGATCTCGGACACGGGACGTGGTCTACACCTGTCGAAGGGACTC

a W L E R V L E A C A L Q P D V D S F P E .

GGAATCCACACTCAATTGGGGAGCAGGGCATGAATCTCTCCGGAGGCCAGAACAGCGG  
 2221 ----- + ----- + ----- + ----- + ----- + 2280  
 CCTTAGGTGTGAAGTTAACCCCTCGTCCCGTACTTAGAGAGGCCCTCGGTCTCGGCC

a G I H T S I G E Q G M N L S G G Q K Q R .

CTGAGCCTGGCCCGGGCTGTATACAGAAAGGCAGCTGTGTACCTGCTGGATGACCCCTG  
 2281 ----- + ----- + ----- + ----- + ----- + 2340  
 GACTCGGACCGGGCCGACATATGTCTTCCGTCGACACATGGACGACCTACTGGGGAC

a L S L A R A V Y R K A A V Y L L D D P L .

GCGGCCCTGGATGCCACGTTGCCAGCATGTCTCAACCAGGTATTGGCCTGGTGG  
 2341 ----- + ----- + ----- + ----- + ----- + 2400  
 CGCCGGGACCTACGGGTGCAACCGGTGTACAGAAGTTGGTCCAGTAACCCGGACCACCC

a A A L D A H V G Q H V F N Q V I G P G G .

CTACTCCAGGGAAACAACACGGATTCTCGTGACGCACGCACCCACATCCTGCCAGGCT  
 2401 ----- + ----- + ----- + ----- + ----- + 2460  
 GATGAGGTCCCTTGTGCTTAAGAGCACTGCGTGTGAGGTGTAGGACGGGTCCGA

a L L Q G T T R I L V T H A L H I L P Q A .

GATTGGATCATAGTGGCAAATGGGCCATCGCAGAGATGGGTTCTACCAAGGAGCTT  
 2461 ----- + ----- + ----- + ----- + ----- + 2520  
 CTAACCTAGTATCACGACCGTTACCCGGTAGCGTGTACCCAAGGATGGTCTCGAA

a D W I I V L A N G A I A E M G S Y Q E L .

CTGCAGAGGAAGGGGGCCCTCGTGTCTGGATCAAGCCAGACAGCCAGGAGATAGA  
 2521 ----- + ----- + ----- + ----- + ----- + 2580

Figure 15F

GACGTCTCCTCCCCGGGAGCACACAGAAGAC~~T~~AGTTGGTCTGTCGGCCTATCT  
 a L Q R K G A L V C L L D Q A R Q P G D R .  
 GGAGAAGGAGAACAGAACCTGGGACCAGCACCAAGGACCCAGAGGCACCTCTGCAGGC  
 2581 -----+-----+-----+-----+-----+-----+ 2640  
 CCTCTCCCTTTGTCTTGGACCCCTGGCGTGGTCTGGGTCTCCGTGGAGACGTCCG  
 a G E G E T E P G T S T K D P R G T S A G .  
 AGGAGGCCGAGCTTAGACGCGAGAGGTCCATCAAGTCAGTCCCTGAGAAGGACCGTACC  
 2641 -----+-----+-----+-----+-----+-----+ 2700  
 TCCTCCGGGCTCGAATCTCGCTCTCCAGGTAGTTAGTCAGTCAGGGACTTTCCCTGGCATGG  
 a R R P E L R R E R S I K S V P E K D R T .  
 ACTTCAGAACAGCCCAGACAGAGGTTCTCTGGATGACCTGACAGGGCAGGATGGCCAGCA  
 2701 -----+-----+-----+-----+-----+-----+ 2760  
 TGAAGTCTCGGGCTGTCTCCAAGGAGACCTACTGGGACTGTCCCGTCTACCGGTCGT  
 a T S E A Q T E V P L D D P D R A G W P A .  
 GGAAAGGACAGCATCCAATACGGCAGGGTGAAGGCCACAGTCACCTGGCCTACCTGCGT  
 2761 -----+-----+-----+-----+-----+-----+ 2820  
 CCTTTCCTGTCGTAGGTTATGCCGTCCACTTCCGGTGTACGTGGACCGGATGGACGCA  
 a G K D S I Q Y G R V K A T V H L A Y L R .  
 GCCGTGGGCACCCCCCTCTGCCTCTACGCACTCTCCTCTGCCAGCAAGTGGCC  
 2821 -----+-----+-----+-----+-----+-----+ 2880  
 CGGCACCCGTGGGGGAGACGGAGATGCGTGAGAAGGAGAAGGAGACGGTCCTACCGG  
 a A V G T P L C L Y A L F L F L C Q Q V A .  
 TCCTTCTGCCGGGCTACTGGCTGAGCCTGTGGCGGACGACCCCTGCAGTAGGTGGCAG  
 2881 -----+-----+-----+-----+-----+-----+ 2940  
 AGGAAGACGGCCCCGATGACCGACTCGGACACCCGCCTGCTGGACGTATCCACCGTC  
 a S F C R G Y W L S L W A D D P A V G G Q .  
 CAGACGCAGGCAGCCCTGCGTGGCGGATCTCCTGGCTCTCGGCTGTCTCCAAGGCCATT  
 2941 -----+-----+-----+-----+-----+-----+ 3000  
 GTCTGCGTCCGTGGGACGCACCGCCCTAGAAGCCGAGGAGCCGACAGAGGTTGGTAA

Figure 15G

a Q T Q A A L R G G I F G L L G C L Q A I .

GGGCTGTTGCCTCCATGGCTGCGGTGCTCCTAGGTGGGGCCCGGGCATCCAGGTTGCTC  
3001 -----+-----+-----+-----+-----+-----+ 3060  
CCCGACAAACGGAGGTACCGACGCCACGAGGATCCACCCGGGCCGTAGGTCCAACGAG

a G L F A S M A A V L L G G A R A S R L L .

TTCCAGAGGCTCTGTGGGATGTGGTGCATCTCCATCAGCTTCTTGAGCGGACACCC  
3061 -----+-----+-----+-----+-----+ 3120  
AAGGTCTCGAGGGACACCCCTACACCACGCTAGAGGGTAGTCGAAGAAACTCGCCTGTGGG

a F Q R L L W D V V R S P I S F F E R T P .

ATTGGTCACCTGCTAAACCGCTTCTCCAAGGAGACAGACACGGTTGACGTGGACATTCCA  
3121 -----+-----+-----+-----+-----+ 3180  
TAACCAGTGGACGATTGGCGAAGAGGTTCTCTGTCTGTGCCAAGTCACCTGTAAGGT

a I G H L L N R F S K E T D T V D V D I P .

GACAAACTCCGGTCCCTGCTGATGTACGCCCTGGACTCCTGGAGGTAGCCTGGTGGTG  
3181 -----+-----+-----+-----+-----+ 3240  
CTGTTGAGGCCAGGGACGACTACATGCGAAACCTGAGGACCTCCAGTCGGACCACAC

a D K L R S L L M Y A F G L L E V S L V V .

GCAGTGGCTACCCACTGGCCACTGTGGCCATCCTGCCACTGTTCTCTACGCTGGG  
3241 -----+-----+-----+-----+-----+ 3300  
CGTCACCGATGGGTGACCGGTAGGACGGTAGACAAAGAGGAGATGCGACCC

a A V A T P L A T V A I L P L F L L Y A G .

TTTCAGAGCCTGTATGTGGTTAGCTCATGCCAGCTGAGACGCTGGAGTCAGCCAGCTAC  
3301 -----+-----+-----+-----+-----+ 3360  
AAAGTCTCGGACATACACCAATCGAGTACGGTCACTCTCGAACCTCAGTCGGTCATG

a F Q S L Y V V V S S C Q L R R L E S A S Y .

TCGTCTGCTGCTCCCACATGGCTGAGACGTTCCAGGGCAGCACAGTGGTCCGGCATTG  
3361 -----+-----+-----+-----+-----+ 3420  
AGCAGACAGACGAGGGTGTACCGACTCTGCAAGGTCCCGTGTCAACCAGGCCGTAAAG

Figure 15H

a S S V C S H M A E T F Q G S T V V R A F .

CGAACCCAGGCCCTCTGGCTCAGAACATGCTCGCTAGATGAAAGCCAGAGGATC  
 3421 -----+-----+-----+-----+-----+-----+ 3480  
 GCTTGGGTCCGGGGAGAACACCGAGTCTTGTACGAGCGCATCTACTTCGGTCTCCTAG

a R T Q A P L V A Q N N A R V D E S Q R I .

AGTTTCCCAGACTGGTGGCTGACAGGTGGCTGGCCAATGTGGAGCTCCTGGGAAT  
 3481 -----+-----+-----+-----+-----+-----+ 3540  
 TCAAAGGGCGCTGACCACCGACTGTCCACCGAACGCCGGTACACCTCGAGGACCCCTTA

a S F P R L V A D R W L A A N V E L L G N .

GGCCTGGTGTTCAGCTGCCACGTGTGCTGAGCAAAGCCCACCTCAGTGCTGGC  
 3541 -----+-----+-----+-----+-----+-----+ 3600  
 CGGACCAACAAACGTCGACGGTGACACGACACGACTCGTTGGAGTCACGACCG

a G L V F A A A T C A V L S K A H L S A G .

CTCGTGGCTTCTGTCTGCTGCCCTCCAGGTGACCCAGGCAGTCAGTGGTTGTT  
 3601 -----+-----+-----+-----+-----+-----+ 3660  
 GAGCACCCGAAGAGACAGAGACGACGGGAGGTCCACTGGTCCGTACGTACCCAAACAA

a L V G F S V S A A L Q V T Q A L Q W V V .

CGCAACTGGACAGACCTAGAGAACACGATCGTGTCACTGGAGCGGATGCAGGACTATGCC  
 3661 -----+-----+-----+-----+-----+-----+ 3720  
 GCGTTGACCTGTCTGGATCTTGTCTAGCACAGTCACCTCGCCTACGTCTGATACGG

a R N W T D L E N S I V S V E R M Q D Y A .

TGGACGCCAAGGAGGCTCCCTGGAGGCTGCCACATGTGAGCTCAGCCCCCTGGCCT  
 3721 -----+-----+-----+-----+-----+-----+ 3780  
 ACCTGCGGGTCTCCGAGGGACCTCCGACGGGTGTACACGTCGAGTCGGGGGACCGGA

a W T P K E A P W R L P T C A A Q P P W P .

CAGGGCGGGCAGATCGAGTCCGGACTTGGCTAAGATACCGACCTGAGCTCCGCTG  
 3781 -----+-----+-----+-----+-----+-----+ 3840  
 GTCCCGCCCGTCTAGCTAAGGCCCTGAAACCCGATTCTATGGCTGGACTCGAGGGCGAC

a Q G G Q I E F R D F G L R Y R P E L P L .

Figure 15I

3841 GCTGTGCAGGGCGTGTCCCTCAAGATCCACGCAGGAGAGAAGGTGGCATGTTGGCAGG  
 3841 ----- + ----- + ----- + ----- + ----- + 3900  
 CGACACGTCCCCGACAGGGAGTTAGGTGCGTCCTCTTCCACCCGTAGCAACCGTCC

<sup>a</sup> A V Q G V S L K I H A G E K V G I V G R .

3901 ACCGGGGCAGGGAAAGTCCTCCCTGGCCAGTGGGCTGCTGCGGCTCCAGGAGGCAGCTGAG  
 TGGCCCCGTCCCTTCAGGAGGGACCGGTACCCGACGACGCCAGGTCCCTCCGTCGACTC

a T G A G K S S L A S G L L R L Q E A A E -

GGTGGGATCTGGATCGACGGGTCCCCATTGCCACGTGGGCTGCACACACTGCGCTCC  
 3961 — + — + — + — + — + 4020  
 CCACCCCTAGACCTAGCTGCCAGGGTAACTGGGTGCAACCCGACGTGTGTGACGCCAC

a G G I W I D G V P I A H V G L H T I R S -

```

AGGATCAGCATCATCCCCCAGGACCCATCCTGTTCCCTGGCTCTGCAGGATGAACTC
4021 + + + + + + + + 4080
TCCTAGTCGTAGTAGGGGGTCTGGGTAGGACAAGGGACCGAGAGACGCCACTTGGAG

```

a R I S I I P Q D P I L F P G S I R M N I -

4081 GACCTGCTGCAGGAGCACTGGACGAGGCTATCTGGCAGCCCTGGAGACGGTGAGCTC  
 4140 CTGGACGACGTCCTCGTAGCCTGCTCCGATAGACCCGTCGGGACCTCTGCCACGTCGAG

a DLLQEHSDEAIWAALLETVOI -

AAAGCTTGGTGGCCAGCCTGCCGGCCAGCTGCAGTACAAGTGTGCTGACCGAGGCGAG  
 4141 — + — + — + — + — + 4200  
 TTTCGGAACCACCGGTGGACGGGCCGGTCAGCTCATGTTCACACGACTGGCTCCGCTC

a K A L V A S L P G Q L Q Y K C A D R G E

4201 GACCTGAGCGTGGGCCAGAACAGCTCCTGTGTCTGGCACGTGCCCTCTCCGGAAAGACC  
 4260 CTGGACTCGCACCCGGTCTTGTGAGGACACAGACCGTGCACGGGAAGAGGGCTCTGG

a D L S V G Q K Q L L C I A R A I I B K T

**Figure 15J**

4261 CAGATCCTCATCCTGGACGAGGCTACTGCTGCCGTGGACCCCTGGCACGGAGCTGCAGATG  
 4320 GTCTAGGAGTAGGACCTGCTCCGATGACGACGGCACCTGGGACCGTGCCTCGACGTCTAC

a Q I L I L D E A T A A V D P G T E L O M

4321 CAGGCCATGCTCGGGAGCTGGTTGCACAGTGCACGTGCTGCTCATTGCCAACCGCCTG  
 4380 GTCCGGTACGAGCCCTGACCAAACGTGTCACGTGACACGACGAGTAACGGGTGGCGGAC

<sup>a</sup> Q A M L G S W F A Q C T V I I I A H R I

4381 CGCTCCGTATGGACTGTGCCGGGTTCTGGTCATGGACAAGGGGCAGGTGGCAGAGAGC  
 4440 GCGAGGGCACTACCTGACACGGGCCAAGACCAGTACCTGTTCCCCGTCCACCGCTCTCGC

<sup>a</sup> R S V M P C A B V I V M P K S O V A S

GGCAGCCGGCCAGCTGCTGGCCAGAAGGGCTGTTTACAGACTGGCCAGGAGTCA  
 4441 — + — + — + — + — + 4500  
 CCGTCGGGCGGGTCGACGACCGGGCTTCCCGGACAAAATGCTGACCCGCTGACT

<sup>a</sup> G S P A Q I I A Q K G + E Y R I A C 5 3

GGCCTGGTCTGA  
4501 — + — 4512  
CCGGACCAAGACT

a G L V i .

**Figure 15K**